

ELEMENTS OF ECONOMIC THEORY

BY

ERICH ROLL

PROFESSOR OF ECONOMICS AND COMMERCE
IN THE
UNIVERSITY COLLEGE OF HULL

OXFORD UNIVERSITY PRESS
LONDON : GEOFFREY CUMBERLEGE

OXFORD UNIVERSITY PRESS
AMEN HOUSE, E.C. 4
London Edinburgh Glasgow New York
Toronto Melbourne Capetown Bombay
Calcutta Madras
GEOFFREY CUMBERLEGE
PUBLISHER TO THE UNIVERSITY

FIRST EDITION 1937
REPRINTED 1938, 1946

PRINTED IN GREAT BRITAIN

PREFACE

THIS analysis of the elementary principles of Economics is largely based on a course of lectures which I have been giving for some years to Intermediate students. It is written primarily as a text-book for students in their first year at the university; but I hope that those in extra-mural classes, those preparing for professional examinations, pupils studying for the Higher School Certificate and, also, some general readers may find it useful.

In a book of elementary exposition, controversy is, of necessity, pushed into the background and there appears a larger measure of agreement than exists in reality. This is particularly so in my formulation of the theory of value, in which the traditional marginal utility approach is preserved. The picture which results is, perforce, drawn on 'orthodox' lines. But it must not be thought that I am unaware of, or necessarily unsympathetic to, the many criticisms which are directed not only against points of detail but also against major aspects and, even, against the whole approach of the kind of Economics here presented. Some of the minor controversial issues I have tried to indicate. Here and there, I have also endeavoured to point out the problems which, at a more advanced stage, should engage the attention of the student. And while, in the main, my task has been to present in broad outline Economic Science as currently

taught, I have made an attempt, at the end of this book, to show the character of different approaches and their implications—particularly in relation to policy.

The technique used is the common property of all economists; all experts will be familiar with the sources. I have therefore omitted references which would only have been burdensome to the student. For the benefit of those who wish for guidance in further reading, I have added a short bibliography.

I wish to acknowledge my indebtedness to my wife and to my friend Dr. J. Bronowski who have eliminated innumerable blemishes from these pages; to two readers whose criticism has enabled me to remove many obscurities and ambiguities; and to my father for his constant help and advice.

E. R.

January, 1937

CONTENTS

PART ONE. THE ECONOMIC PRINCIPLE AND THE ECONOMIC WORLD

1. The Problem of Definition	1
2. A definition of Economics	9
3. The Scope of Economics	13
4. The Economic Structure	17
5. The Method of Economics	29

PART TWO. THE WORKING OF THE ECONOMIC PRINCIPLE

I. THE THEORY OF CONSUMPTION	38
1. A division of Economic Theory	38
2. Wants	42
3. Goods	48
4. Utility	51
5. The Maximum of Satisfaction	57
II. THE THEORY OF EXCHANGE	64
1. The Conditions of Exchange	64
2. Isolated Exchange and Monopoly	67
3. Bilateral Competition	76
4. Monopoly. Perfect and Imperfect Competition	81
5. Money	87
6. The Communal Scale	91
7. The laws of Exchange	100
8. Complements and Substitutes	106
9. Equilibrium	110
III. THE THEORY OF PRODUCTION	115
1. The Scope of the Theory	115
2. Production in the isolated economy	119
3. Production in the exchange economy	122
4. Single-factor production.	126

5. Producer's Surplus	135
6. Joint Products	143
7. Complex Production	149
8. Rent	166
9. Wages	169
10. Capitalistic Production	179
11. Profits	195
PART THREE. PROBLEMS OF CHANGE	
1. The quality of changes in data	202
2. Population	210
3. Tastes	217
4. Technical Progress	227
5. The problem of Policy	237
6. The Trade Cycle	247
CONCLUSION	262
BIBLIOGRAPHY	269
INDEX	273

PART ONE
THE ECONOMIC PRINCIPLE AND
THE ECONOMIC WORLD

I. THE PROBLEM OF DEFINITION

Uses of the word 'economic'. The term 'economic' is so freely used to-day that to most people it seems to possess a fairly well-marked meaning. We read, for example, that a local authority has rejected a scheme for the erection of a bridge on the ground that it is not an economic proposition. We know what this means: much as the bridge might have added to the amenities of the district, much, even, as it might have benefited trade by reducing distance and transport costs, it was felt that the cost of it was excessive in relation to the benefits to be derived from it. Thus, in this use of the term there is involved a comparison between cost and benefit.

But this is not always the meaning which it carries. We know that one of the methods of expressing disapproval with a country's policy is for other countries to impose 'economic sanctions', to enforce an economic boycott. Here, what is involved is a refusal to exchange with a country, regarded as an international wrong-doer, the goods and services normally exchanged with it; to cease buying and selling, lending and borrowing. In this use of the word 'economic' there is, then, implied a reference to trade, the buying and selling of goods and services.

Yet again, we find the word being used if it is desired to distinguish between different sets of arguments. Thus, the politician on his election platform might state that while on economic grounds the growing of wheat on a very large scale in England was not to be regarded as altogether profitable; yet, because of its importance as a staple food, it was desirable to encourage its production so as to diminish dependence on foreign supplies in time of war. Here, too, there appears to be a balancing of cost and benefit, but it is combined with an appeal to certain desirable objects which make even high cost justifiable.

Again, when an eminent K.C. gives up a practice which is reputed to bring him an income of £30,000 a year in order to become a judge at a salary of £5,000 a year, he is often said to have sacrificed economic interests for public service or the desire for social distinction. And, in the hackneyed phrase that man does not live by bread alone, the higher activities of man appear to be contrasted with his baser economic pursuits.

The need for definition. If economics is the study of economic problems, it is necessary to discover whether in all the instances which have been mentioned the word economic is rightly used. If it is, we have to find what quality is common to them all. It is necessary to define our subject-matter. Definition is not the idle pursuit of the pedant; it is an essential part of any systematic discipline that the limits of the field which it sets out to cultivate should be clearly

marked. Through definition we make certain that we are investigating only problems that are related and all the problems that have the same characteristics. Only by defining our subject-matter can we ensure that we shall know when our conclusions are applicable and when they are not. It is by definition that we assign to each discipline its room in the building of knowledge.

Necessary as definition is, it is by no means easy. Indeed, we shall find that some of the most far-reaching disagreements in economics arise from differences in the conception which different practitioners have of their subject. This is partly due to the youth of economics. The first systematic treatise of it was published only 150 years ago, at a time when most sciences could look back upon centuries of progress. Since then it has made such rapid strides—due, no doubt, to the speed with which new problems have been created for it—that economists have had little time to worry about neat definitions. They have built up a large body of conclusions; and it has been left to modern economists to discover from the nature of those conclusions what economics is really concerned with. Such interpretations are bound to differ. But since the authority of economics is invoked so often to-day—and so often rejected—economists are forced to state clearly what is the scope of their competence. At the end of our study we shall have to return to this problem of different definitions; for the moment we shall have to find some definition which will serve as a satisfactory working basis—containing not all the truth perhaps, but

4 ECONOMIC PRINCIPLE AND ECONOMIC WORLD
sufficient of it to serve as a foundation for our more detailed analysis.

Different kinds of human activity. It is clear that economics deals with human activity. It must be equally obvious that this is too wide a field to cover. Some early economists—and their view is still widely held—tried to distinguish between different kinds of human activity and to assign one of them to economics. What the economist studies, they said, was how people made, distributed, and exchanged the material goods which they needed for their livelihood. The boots we wear, the bacon we eat, the houses we live in: these constitute the material means of our lives; and in economics we study how they are provided and used, why we have more of one thing than another, or more at one time than at another time, or why some people have more of them than other people.

At first sight this certainly seems to cover what most of us mean when we talk about economic problems. It appears to include our example of the bridge and certain aspects of the refusal to buy a country's lemons or sell it coal. But some of our other usages do not fit quite so easily. What, for example, of our self-sacrificing lawyer? Honour and glory are certainly not 'material' goods in the ordinary sense of the word; yet they seem to have been acquired in exchange for all the material things which a larger income could buy—country houses and Rolls-Royce cars, for instance. We often exchange such 'material' benefits for others that are 'non-material'; and it may be

argued that even 'material' goods serve 'non-material' ends: certainly, the enjoyment which we derive from them is always 'non-material'. This is most obviously so with the scholar's books, the painter's canvas, the musician's violin.

Nor are these the only difficulties which the separation of different kinds of activity involves. The early economists argued at great length about what was and what was not productive labour. Clearly, such a distinction was only a corollary of that between 'material' and 'non-material' goods; and if one adopted this separation it was inevitable that one should have to regard as unproductive (as did Adam Smith) 'some both of the gravest and most important and some of the most frivolous professions; churchmen, lawyers, physicians, men of letters of all kinds, players, buffoons, musicians, opera-singers, opera-dancers, &c.' With this definition we should certainly be faced with difficulties, and modern economists have argued that economics would be in danger of becoming a kind of medieval scholasticism.

Different motives of human activity. The earlier economists made yet another attempt to divide the field of human activity: a division according to the motive which inspired it. Self-interest was regarded as the 'economic motive', and was distinguished from all those other motives, of friendship, religious devotion, philanthropy, vindictiveness, and prejudice which inspire human conduct. Sometimes this distinction of motives was formulated in such a way as

6 ECONOMIC PRINCIPLE AND ECONOMIC WORLD

to give rise to the idea that what economics was concerned with was not man, but some strange and unreal being, the 'economic man', a grasping sort of individual who reduced everything to a calculation of money gain and loss. When understood in this way our new definition of the subject-matter of economics could rightly be attacked—as, indeed, it was and still is—on the ground that it dealt with only one side of human nature, while it could easily be shown that human nature is essentially complex. It was argued that in adopting this definition economics was basing itself upon a very primitive psychology, and any advance in the study of the motives of human behaviour would, therefore, upset the conclusions of economics.

Moreover, so narrow a view of economic motive would exclude from our consideration many problems that are usually regarded as falling within the province of the economist; and much of what was said earlier about the 'material' and 'non-material' would apply. What, for example, can we make of the fact that some occupations will attract people more than others, even though monetary remuneration is smaller, because they offer advantages of another kind? Here one would rightly suspect the existence of an economic phenomenon. Yet it is one with which the sort of self-interest on which the fiction of the 'economic man' was based is incapable of grappling. It is not necessary to have much formal knowledge of economics to realize that the amounts of goods and services offered and demanded, and their prices, are proper matters for

economic study. In order to study these offers and demands it is not necessary to assume that they are the results of one motive alone. We shall see that we must sometimes make certain assumptions which appear to include a reference to motive; but for the moment we need not restrict the field of economic study in that way at all.

Wants and their satisfaction. Thus, our two definitions are seen to be too narrow. We want one which will include non-material as well as material objects with which human activity is concerned, and one which will allow for all the motives by which such activity may be inspired. It is quite easy to widen the first of our definitions. We simply have to leave out the word 'material' and say that we are concerned with activity directed towards the satisfaction of wants. Similarly, we should make it clear that we are not inquiring into the motives of that activity, but only into the observed conduct of individuals. We leave it to other sciences, psychology, for example, to discover the motives behind such conduct. Our two fetters have disappeared, and we can say that what we are interested in is the process of want-satisfaction.

But stated in these terms, does not our definition now appear to be too wide? Have we not taken upon ourselves more than we can usefully examine by one science alone? Let us look at some of the things which our definition would oblige us to study. At once a variety of problems present themselves. Why, for

8 ECONOMIC PRINCIPLE AND ECONOMIC WORLD

example, are our wants what they are? As we shall see, many branches of knowledge try to answer this question: the biological sciences for our physical wants; history and social psychology, for those wants which are of a more complicated nature and are determined by the social system in which we live.

Again, the extent to which wants can be satisfied depends largely on geographical and technical conditions and on the degree of scientific knowledge which we possess. For these factors there are other special branches of study. Is there anything left, then, which economics can contribute to these aspects of the process of want-satisfaction?

The question why man wants motor-cars is a very interesting and important one; the biologist, social historian, and philosopher may all be able to suggest answers to it. They may point to the comparative inefficiency of the human body as a means of locomotion, and to the need for mechanical aid. They may even suggest that there is an undefinable urge in human beings to explore all their potentialities, and they may describe the historical process in which this urge has worked itself out. Elaborate theories concerned with the nature and evolution of mankind may be put forward to explain the fact that in the twentieth century so many people experience a want for motor-cars. To the further question, why most men should desire a particular make of car, the psychologist may be able to give an answer. He may explain it by a desire—possibly inherent in man,

possibly due to social circumstances—to achieve distinction over his fellows. Or he may discover an answer in the kind of advertising used by the makers of the particular car, and he may, thus, be led to study the effects on people's minds of suggestions conveyed through advertisements.

Similarly, when we study the other side, the way in which these wants (whatever may be their cause) are satisfied, we find a number of sciences busy contributing to the answer. To the making of a motor-car there goes the knowledge of the mathematician, physicist, and chemist, and the skill of the mining engineer, the metallurgist, and the mechanical and electrical engineers, to mention only a few. The geologist and geographer contribute in the search for materials and factory sites, and the architect in the building of the factory. Thus, in the last analysis, the process of want-satisfaction is studied, explained, and aided by all branches of human knowledge.

2. A DEFINITION OF ECONOMICS

Limited and unlimited resources. What is the special contribution which economics can make? Which aspect of the human activity of want-satisfaction is its peculiar field of study? To arrive at a satisfactory answer, let us consider a very simple example of human activity directed towards satisfying a want. Air is one of our most urgent wants. We need a supply of it if our very existence is to continue. The biologist will tell us why our organism needs air and

how it uses it; while the chemist and physicist will be able to tell us something about the nature of the supply of air. Generally, the satisfaction of this want presents no problem; in the country, in a healthy climate, it is ensured by an automatic process. But to ensure an adequate supply of air in a large London restaurant is a very real problem indeed. The knowledge and skill of engineers are now needed to install a system of ventilation; the supply of air has become a technical problem. But something else is required: the provision of air now involves an expenditure of time, of energy, of materials; and it is the need for this expenditure which makes it into an economic problem as well. What is it, then, which distinguishes the problem of the provision of air in the second instance from that in the first? It is the fact that in the London restaurant air has ceased to be free: its quantity is no longer unlimited in relation to the need for it.

Thus, the objects which are capable of satisfying our wants may be available in limited or unlimited quantities; and even the same object may, under different conditions, be unlimited or limited in relation to the need for it. There may be communities for which the securing of food presents no problems beyond those of a technical nature—which kind is wholesome and which is not, for example; food is had for the asking. We know only too well that this is not so in the civilized communities of which we have knowledge. It is no longer enough to know where

different kinds of food are to be found, what is their nature and effect, which kinds are more nourishing than others. The quantity of food available, in the form and place and at the time required by the consumer, is now limited in relation to the need for it. This is so partly because of the larger number of people whose appetite has to be satisfied, partly, perhaps, because of the greater appetite of each, and partly because of the more refined nature of that appetite.

Even if it can be shown that some foodstuffs are, absolutely, available in superabundance, yet to transport them from their source of supply to where they are needed and to present them in a form adapted for consumption requires the expenditure of much time and effort; and these are certainly limited in relation to all the uses to which they can be put.

Scarcity. *Scarcity* of the means for satisfying our wants is, thus, the characteristic of an economic problem. The word 'scarcity' is normally used in a more restricted sense than that which we have used. When a book is described as scarce in a bookseller's catalogue we should take that as meaning that only very few copies of this book are available. Radium is scarce in this everyday sense of the word; we should not, however, regard water or air or wheat or motor-cars as being scarce in the same sense. Yet, in economic parlance, scarcity begins as soon as the quantity of a thing ceases to be unlimited. It is a *relative* concept expressing the relation between our wants and

the means for satisfying them. In this sense it is not only a Gutenberg Bible, or a First Folio, which is scarce, but also the modern twopenny thriller. True, the quantity of the former can no longer be increased, while that of the latter can; but, in order to increase it, use will have to be made of paper and printers' ink and machines, all of which are scarce in the economist's sense of the word.

It is very difficult to find many things in the community in which we live which do not exhibit this quality of relative scarcity. And in relation to all the ends which we wish to achieve, our means for achieving them, reduced to their original forms—the resources provided by nature, our physical and mental energy, the time at our disposal—are definitely limited.

Alternative uses. Reduced to these forms, another characteristic of these limited means becomes particularly obvious: they can be used in alternative ways. This is true of nearly all the means of which we dispose. We can drink our milk or have it made into umbrella handles. We can eat our bread and butter or leave it to see it reappear as bread-and-butter pudding. We cannot do both. In these cases the range of possibilities is narrow: two or three will generally exhaust them. But the range of possible uses to which the original constituent elements of, say, bread can be put is very great indeed. There are many things which might be produced (and wants which might be satisfied) out of the wheat and the yeast and the iron and

the coal and the labour which have helped to make it. But, whatever it is which is produced, and whatever the want which is satisfied, the other possibilities of production and the other want-satisfactions are forgone.

Thus, the limitation of our resources and their ability to satisfy different wants create a problem, the problem of choosing how to dispose of these resources: which wants to satisfy, since they cannot all be satisfied. The economic problem is, therefore, essentially a problem arising from the necessity of *choice*—choice of the manner in which limited resources with alternative uses are disposed of. It is the problem of the husbandry of resources.

Economics studies the activity of husbandry; indeed, the literal translation of economy is 'law of husbandry'. So long as our resources (at any rate, in their original form) are scarce in our sense of the term, so long will an economic problem exist; so long, also, will there be room for economics as a special branch of study.

3. THE SCOPE OF ECONOMICS

The quality of choice. The definition we have finally arrived at still leaves us with an almost overwhelmingly large ground to cover. It has, indeed, been argued that choice, which we have taken as the characteristic quality of economic conduct, is inherent in all human conduct. All human activity, it is said, can be explained in terms of acts of preference; the

boundary line between preference and non-preference is the same as that between action and inaction. For as soon as we act, we decide; and a decision involves a choice—a preference of A to B. Such a choice may be purposeful, that is to say, the action may be preceded by deliberation on the alternatives and by a conscious preference of one. Or we may not even wish to go as far as to investigate the reasons for the act, and we may simply infer the existence of preference from the fact that a decision has been made. For this purpose, the word preference may, perhaps, be a little less open to misinterpretation than the word choice. We may simply take it to signify that when an individual does A rather than B, he has ‘preferred’ A to B.

However that may be, one thing is clear: to study the problems of choice means to study an enormous part, if not the whole, of human conduct under that one aspect. Many economists do, indeed, claim this as their competence. They point to the fact that the terms relating to choice used in economics are often applied in everyday usage to what is not commonly regarded as an economic activity. For we speak of cost and price and even capital and interest in relation to many non-trading activities. In the Bible we read of the wages of sin; and dictators may tell us that ‘honour is too high a price to pay for peace’.

Choice in the market. It is, fortunately, not necessary for us, at this stage, to decide whether the claims of these economists are justified or not. For all econo-

mists are agreed that from amongst the mass of phenomena of choice they only study a certain number; some economists say because these are their proper field of study; others say because these are the most interesting and most important of all those which they could study, and because they are phenomena for which their apparatus of analysis will yield the most fruitful results.

We study, then, those acts of choice which habitually occur in the market. Those acts which, more or less regularly, give rise to offers and demands for goods and services and to exchange transactions. It is the habitual nature of these acts which makes them susceptible to systematic analysis. Moreover, it is these market transactions which obtain a quantitative appearance and become, thus, capable of measurement—again one of the preconditions of systematic study. This idea is sometimes conveyed by saying that we study only those acts of choice which are habitually part of the ‘pricing’ process, i.e. which become capable of expression in terms of money.

To take a walk in one direction rather than another, or to have an afternoon’s nap rather than to take a walk, involves choice. But it is not an act of choice that can be expressed in terms of money. The decision whether to take a walk or to go to the cinema, or the doctor’s decision whether to have an afternoon in the country or to be at home and available to patients, already involve some monetary calculation. There are many other decisions of this kind that have to be

made regularly, and involve monetary reckoning. Whether to buy this or that, whether to sell more or less at a certain price, whether to hire out one's labour under certain conditions (including a monetary reward), whether to buy one's want-satisfaction now or in the future—all these are the acts of choice which economics can most fruitfully study.

Choice and the social system. There is yet another sense in which our definition of economics as the study of the disposal of limited resources with alternative uses can be said to be of universal application. The necessity of choice is independent of the social system in which it takes place. It was said above that the economic problem, according to our definition, would remain so long as scarcity remained. Now this scarcity could, and does, exist in many different societies. It exists in the one-man community of Robinson Crusoe, in the patriarchal tribe of Central Africa, in medieval, feudal Europe, in modern, capitalist America, and in communist Russia. It is likely to continue certainly in our generation, if not, as some economists have emphasized, so long as we remain this side of Paradise.

Now in so far as we can develop our original definition of economics, the conclusions which we can derive from it will be of as universal a character as the original definition itself. They will be applicable to all societies, and will thus become very similar to the generalizations of the physical sciences. The laws of choice, like the laws of gravitation, will be independent of

the legal and political framework in which they work.

We do possess a body of generalizations of such universal validity. The first sections of the next part of this book will be devoted to their exposition. But we shall find that in order to give these generalizations some measure of concreteness, in order to make them tell us more about reality, we shall, at a very early stage, have to make certain assumptions about the social framework in which they work. We shall accordingly discover that below the highest level of abstraction (that which yields conclusions of the most general quality), our analysis and our conclusions are determined by the social system which we are examining.

We cannot pursue this argument any farther at this point—its significance can only be assessed when we have seen the apparatus of economic analysis at work. For now it must suffice to state that we have here a further limitation upon our universal definition, one set by the social framework within which economic activity is carried on.

4. THE ECONOMIC STRUCTURE

The division of labour. It is necessary to outline some of the fundamental facts of our present social system in so far as they affect economic activity, in order that a background of some degree of realism may be present from the very beginning. In the following few pages devoted to this exposition we are

not descending to a very low level of abstraction; we are still outlining such institutional qualities of our social system as have been present, in more or less developed form, for the last hundred years or more. We are, that is to say, not describing any of the specific traits of the social structure of, say, England of the present day. But rather are we endeavouring to give a rapid survey of the general features of the modern industrial system, some of which were undoubtedly present before the modern system was fully developed, others of which may already be disappearing.

What probably strikes one most at the first attempt to reduce our industrial system to a few general characteristics is the extent to which each individual is dependent upon that system for the satisfaction of his wants. Unlike Robinson Crusoe, we are not only dependent on natural conditions (including our own abilities) to determine the degree to which we can procure our means of livelihood in the widest sense, but also on our fellows in all parts of the world. We are, in fact, not self-supporting; in the vast majority of instances, the members of our community do not make their own clothes and their own dwellings and procure their own food, but they have to rely on getting these things from the sum total of means provided by the community as a whole.

If we survey the mass of material goods which a member of a civilized Western community consumes, we are soon aware of the astonishingly small degree in which he has directly contributed to their provision.

The clothes he wears have been provided by the labour of hundreds of people in different parts of the world: the wool growers in Australia and the cotton growers in Egypt; the people engaged in transporting the raw material; the merchants in London or Liverpool who sell it; the spinners of the yarns and the makers of the wool tops; the weavers, the dyers, and the finishers; the makers of the sewing cotton and of the buttons; the cutters and the tailors—to mention only a few links in a long chain.

The absence of individual self-sufficiency is due to *division of labour*. This process did not by any means start with modern industry; it is to be observed even among very primitive tribes, where shepherds and hunters and craftsmen constitute, as a rule, distinct classes. But during the last two hundred years this division of labour has vastly increased. It has also changed its character. One of the greatest changes in history was the specialization of every individual on the production of a single commodity or service; this led to the era of handicraft. But, to-day, specialization has gone even farther and the tasks involved in the production of a single article are split up and subdivided among many individuals, as in the famous example of the eighteen different operations involved in the making of a pin, cited by Adam Smith. Each man contributes only a small part; and, as has been pointed out by one great economist, few people can now say, 'I made this'.

Machine production. This further development of

the division of labour, some of the results of which we shall study later, was accompanied, and largely made possible, by the enormous increase in the use of mechanical power and machinery. In essence, this is again no new development, for even at very primitive stages of human development man had equipped himself with various tools and instruments in order to aid himself in his attacks upon nature. But the large-scale employment of inanimate power—water, steam, electricity—and the use of highly complicated mechanical contrivances in place of many of the actions of the human body, is more characteristic of our own age than of any that has preceded it.

It means that our want-satisfaction is indirect, not only in so far as we are dependent upon the actions of others, but also in the sense that we have to establish, maintain, and increase a large mechanical equipment, by the aid of which we then procure for ourselves the things which we need. As we shall see again later, this indirect process of want-satisfaction is known as the *capitalistic* method. In this light, the capitalistic process (production with the aid of capital equipment) appears, however, not so typical of modern industry as it is often taken to be. For even the javelin of the savage hunter might, according to this definition, be regarded as capital equipment, and the provision of food, with its help, might be said to be made by capitalistic methods. It is necessary to add further characteristics before a picture typical of none but our own age emerges.

Private property and exchange. Division of labour (and the use of machinery) give rise to two important social institutions, *private property* and *exchange*. Whatever may be the legal definition of private property, it is enough, for our purpose, to regard it as the right to dispose of material objects. This right has always existed for some objects. Articles of consumption, for example, cannot be used without being private property, for in their case use and ownership are identical. The position is a little less clear in regard to durable articles of consumption, for while the benefits of the ownership of bread are exhausted in a single use, those in a necktie are not. In the latter case, therefore, property and use need not be identical. Yet even with regard to durable consumption goods, there has always been a tendency for private property to exist.

The real difference arises in regard to those articles which serve productive purposes, and it is in their case that division of labour has led to the establishment of private property. The benefits of means of production are not exhausted in a single use, nor need the means be owned by those who use them. When, however, specialization begins, it becomes convenient for the tools of production to be adapted to the individuals using them whose property they then become. The handicraft era is characterized by the existence of private property in the primitive means of production then in use. So long, however, as these tools still produce only for the self-sufficiency of the individual (or

his family) there is no great difference between this form of private property and that in consumable goods; it is still equivalent to use.

Modern machine equipment, however, need no longer be the property of those using it, and private property assumes, therefore, a new and impersonal character. This is, moreover, intensified by the development of exchange. In its simplest form exchange is a bartering of the surplus products of the specialists. In its more advanced form, under modern conditions of machine production, it involves the habitual production for the market, i.e. for purposes of exchange. Private property in means of production becomes, then, the administration, by individuals, of these means in a manner dictated by the market.

The market. Exchange, resulting from the progress of division of labour, separates producer and consumer, and producer and product. The isolated and self-sufficient individual is his own master. So, in a sense, is the self-contained family or patriarchal tribe. There is a single will which determines what each member is to do, what is to be produced, how the product is to be shared and consumed. But when the individual ceases to be self-sufficient, when he ceases even to produce a single whole thing, a new impersonal guiding authority appears which determines what he is to produce and how he is to administer his resources.

The mechanism which unites these scattered individuals, and which supplies a new regulator in place

of the old central authority or authorities that have existed right up to modern days, is the *market*. Through production for the market and through the act of exchange, economic activity is made into that social process which division of labour necessitates. Only by bringing his products to the market and by exchanging them with the products of others can an individual satisfy his wants; and his success in the market will determine the degree to which his wants can be satisfied. No conscious regulation takes place, and each individual is still left to choose among the possible alternative ways of using his resources. But his choices will now have to take into account the conditions obtaining in the market; and their effectiveness will depend on the extent to which they are endorsed by the market.

The history of the Western world in the last few hundred years has been marked by a great extension of the sphere of the market. One by one, the medieval barriers have fallen and more and more commodities have been drawn into the circle of exchange. This process has been aided, if not made possible, by the development of a generally accepted medium of exchange, money, which overcomes the inconveniences of barter and enables ratios of exchange to be more precisely expressed. Monetary calculation has become the characteristic feature of our economic system. More and more activities have been made subject to it and have been rationalized by it.

It is through the growth of monetary calculation

that the concept of the 'economic man' has arisen. We have already seen that it is not necessary to assume that those acting in the market are actuated by a single motive. For, on whatever motive an individual acts (and thereby himself influences the market), he is choosing from among the possibilities which the market offers: he is adopting one and rejecting others. But as an assumption that individuals will desire to make the best use of their resources, and that this desire will show itself in their responses to the monetary stimuli of the market, the concept of the 'economic man' is, indeed, one of the important institutional data of the social framework in which we live. As such, it has often been the object of attack by moralists who have objected to the all-pervasiveness of the monetary calculus. And it should be pointed out that, in describing what he regards as an important part of 'things as they are', the economist does not necessarily justify it. It may, however, be admitted that, even from an economist's point of view, some difficulties remain hidden in this concept of the 'economic man'.

The wage-earner. One of the most important of the things which have been drawn into the ever-widening circle of exchange is human labour power. This process involves, on the one hand, legal freedom for the labouring classes, the abolition of slavery and serfdom, and, on the other hand, the organization of production in such a way as to lead to the need for *hired labour*. The former of these developments was accomplished by a struggle lasting many centuries,

which produced increasing degrees of personal freedom in a long evolution from 'status' to 'contract'. The latter development was due both to the progress of division of labour and machine production, and to the concentration of capital and the development of private enterprise.

Private enterprise. Thus, our earlier statement that the modern era is characterized by production for the market must be supplemented by saying that the initiation of this production is in the hands of private individuals who possess, or can obtain, the large resources necessary for modern machine industry. It is the private *entrepreneur* who buys (in the market) labour power, raw materials, and machines, combines them in the process of production, and sells their products in the market. It is assumed that in so doing he is acting as an 'economic man', that is to say, that he will so order his purchases and his sales that he will obtain as large a return for himself as possible. Or, as it is generally formulated, we assume the entrepreneur to be actuated by the profit incentive; and this must, indeed, be regarded as one of the important institutional data of our system. The resources necessary for the initiation of production may be provided either by the entrepreneur himself (who may possess them through inheritance or through his own accumulation), or he may have them provided by other owners of these means through the specialized machinery which has been developed. The existence of credit and of specialized institutions to administer

26 ECONOMIC PRINCIPLE AND ECONOMIC WORLD
it (the banking system, the capital market) is, thus, another characteristic of our society.

Inequality. We have already touched on others; the existence of *inheritance*, the process of *acquisition*, and, allied to them, the existence of *inequalities* in wealth. The development of the economy of the market, money and credit, and the abolition of the privileges of status took place during a long and painful historical process. Our society did not start from 'scratch' but with an enormous heritage of established rights and inequalities of wealth and social position. These must be kept in mind as important institutional facts when we analyse the working of the economic principle of the administration of scarce means in our society. These inequalities tend to be perpetuated by inheritance, restricted though this may be by redistributive taxation; and there remains much in our social environment that tends to create privileged 'strategic' positions of birth, wealth, education or other, often fortuitous, factors.

Inequalities of wealth tend, moreover, to be constantly recreated by the fact that the system encourages an urge towards acquisition of means as a reserve against the vagaries of the market. Acquisition is also one of the most important ways and, often, the only way, to achieve a position of direct influence over the process of production. This urge becomes, therefore, one of the most important social data which we have to take into account.

Competition and Mobility. The existence of the

market has a further result. It leads to the recognition of the existence of *competition*. This implies that the responsiveness to changes in the market situation includes some regard for the action of others who are in the same position as oneself. It means that there are many sellers (of everything saleable, including labour power), and that each will respond to the monetary stimuli of the market in such a way as to secure the best bargain for himself, in competition with his fellow-sellers, and that the same applies to all the buyers. We shall discuss this point in detail later, and we shall see that economic theory analyses situations of varying degrees of competition (including those with a complete absence of it). But the presence of a general tendency towards competition as a broad feature of our economic system has for a long time been one of the institutional assumptions of the body of economic doctrine.

As a corollary of market competition, a high degree of mobility of the different economic elements has been achieved; but only through the drawing of the whole world into the circle of exchange and through the break-down of many of those barriers to mobility which were due to legal and political conditions or to traditional modes of thought. Economists have thus been able to assume that, with exceptions which are known and are the subject of special study, resources will tend to move into and out of different channels of use according to the state of the markets; in other words, that in our society, goods, including labour,

will tend as far as possible to leave those channels in which they obtain lower prices for those in which higher prices are ruling.

Finally, we have to take into account the existence of a general system of laws and political institutions designed to maintain and strengthen the tendencies which we have enumerated. In democratic communities, the last 150 years (excepting the most recent past) have seen a gradual decline of statutory privileges, a gradual increase of political equality, the strengthening of the laws upholding private property, freedom of contract, individual enterprise and the competitive spirit, and the growth of economic machinery and of social arrangements designed to foster mobility.

It is true that legal enactments exist restricting some rights of property and designed, in the interest of certain members of society, to mitigate the forces of competition. It is necessary, at a later stage, to study their nature and significance; but, as a general rule, the working of the economic principle has hitherto been studied within the political and social framework of the liberal society which developed in the nineteenth century.

Acquaintance with recent economic and political developments must show that much of the above description appears fast to be becoming unrealistic. For example, competition appears to be more and more restricted, sometimes with the definite aid of legislation; and the barriers to the movement of goods

and to the migration of labour appear to have left little of the mobility which was mentioned above.

There can be no doubt that important changes are taking place in our institutions; and many of the data which give to the conclusions of economic theory their concreteness and relevance to reality are no longer those which earlier economists had to take into account. These recent developments and, indeed, the general types of changes with which we have to concern ourselves, will form a part of our subsequent inquiries.

But as a first approximation to the facts of the real world much of our description still remains true; and provided we are aware that important modifications will have to be made later, we can proceed to a study of economic theory with that general picture as our background. For whatever changes are taking place and may yet take place in our society, they will first have to make themselves felt as alterations in those elements which constitute our present social framework.

5. THE METHOD OF ECONOMICS

Is economics a science ? Before we can proceed to a discussion of economic theory proper we must examine a few problems concerning the *method* of economic inquiry. In the first place, there has, for a long time, been some controversy over the question whether economics is a science. Much of this controversy is somewhat pedantic and barren. But it is

possible to come to an affirmative conclusion, with due qualifications, on the grounds set out below.

Impartiality. We can say that there are two factors which determine whether a subject of study is or is not a science: the one subjective and the other objective. Scientific study implies, first of all, a certain state of mind of the inquirer. It demands of him the scientific spirit, in other words, the quality of intellectual curiosity, the desire for knowledge of the truth, the ability of sustained intellectual argument. Above all, the quality of disinterestedness is essential. Absence of partiality is the most important qualification of the scientific mind. It does not mean that the scientist must be an inhuman creature devoid of passions and prejudices. But it does mean that, in so far as he is concerned with his own subject of study, he should rule out his own predilections; where he cannot or does not wish to achieve impartiality he should be aware of this fact and explicitly state it.

In the physical sciences such impartiality is, generally, not difficult to attain. This does not mean that research in these sciences is not bound up with the social system which it serves, or that it is not, in the widest sense, practical. But the history of the discovery of some of the most fundamental physical laws shows that bias has generally been absent. After all, the study of logarithms or of the law of gravitation is little affected by the student's views on, say, the institution of private property or the more distant moral effects of monetary reckoning; even though social

historians may be able to prove that the incentives which first led to these studies are to be looked for in social conditions.

As we have defined it, the scientific spirit is obviously more difficult to achieve in economics than in the natural sciences. For economics is, as we have seen, concerned with human behaviour, with the satisfaction of human wants with the aid of limited means, and with the institutions within which this want-satisfaction takes place. In such a study, the inquirer's own conceptions of what is desirable must inevitably intrude upon his selection and handling of the material of his inquiry.

Yet there is no reason why an attempt should not be made to exclude that which is simply desire or prejudice. Here, even less than in the natural sciences, can the student hope to be completely impartial. But he can, at any rate, strive to separate his unbiased analysis of what *is* from his own views of what *ought* to be. In our body of theory, the first, but not the second, auxiliary verb ought to appear. We should carefully distinguish between our *positive* and our *normative* statements. There is no reason why economists should not be as partisan as other people. But they must be constantly on their guard to rule out normative elements from their theoretical analysis; and to state explicitly when they are including them or when they are drawing conclusions from their positive theory which refer to what they regard as desirable courses of action.

32 ECONOMIC PRINCIPLE AND ECONOMIC WORLD

At the end of our inquiries we shall have to take up again the problem of partiality, but, in so far as economics can achieve the above-mentioned separation, it will fulfil the conditions of a science as far as the subjective factor is concerned. It should, however, be pointed out that in so far as the conclusions of economics are applied to concrete problems and are used as guides for action, we are dealing with an art.

The making of laws. We now come to an even more important qualification of a science—an objective one. This concerns the quality of the material studied. How far is this material capable of yielding universally valid generalizations or *laws*? How far are the statements contained in these laws capable of quantitative expressions?

We know that the material of the natural, or physical, sciences possesses these qualities to a very high degree. It is for this reason that, both in English and in French, though not in German, science normally refers to these subjects only. (The German word *Wissenschaft* is much wider in meaning than our word 'science'; it includes any subject susceptible to systematic study.) The natural and physical sciences have developed a large body of laws of universal validity; they are constantly concerned with the quantitative measurement of their material, and with the discovery of relationships that can be expressed quantitatively.

The work of these sciences proceeds by three closely related methods. There is, first, the activity of the experimentalist who observes his material under

specific surrounding circumstances of, for example, temperature, atmospheric pressure, and friction, which he is able to create in his laboratory. He is thus able to show up the characteristics possessed by his material and to arrive at conclusions directly supported by the evidence of experiment.

A second activity is that of re-examining, measuring, and classifying the phenomena which are already known. Finally, empirical results are used as the raw material for the making of laws. To the evidence which experiment has provided, hypotheses are added and generalizations are made. The resulting theories are then again submitted to the test of experiment.

These three procedures are intimately related to one another and, in them, induction and deduction interpenetrate. Little experimental work is done without some preceding hypothesis and *a priori* reasoning. On the other hand, the process of law-making demands the constant check of empirical research.

The laws of these sciences are universally valid in the sense that, if they are correct generalizations, the results will always be the same if the accompanying circumstances are the same. From these universal ideal laws the sciences are, then, able to make concrete statements of a very high degree of probable truth by making allowance for the variation of the 'real' conditions from the 'ideal' ones.

Now, while hypothesis plus verification through the observation of real conditions is open to the economist, he has no laboratory in which he can experiment.

The material of the real world is constantly shifting under his eyes and he is not in a position to manipulate it. Nevertheless, some schools of economics, for example, the historical schools, have suggested that even though the economist cannot experiment, he can observe, he can collect the data of his observations, and he can proceed to discover from them whether they exhibit certain common and recurrent traits.

The difficulty with purely empirical generalization in economics is that it can never yield universally valid results. Whatever the conclusions we arrive at by it, they are inevitably limited by their time and place and other accompanying circumstances. These circumstances are numerous, and their interrelation is extremely complex. It is, therefore, generally impossible to say with certainty which of the many possible causes has, in fact, produced a given result. Nor is it likely that quite the same combination of historical circumstances will ever repeat itself.

Much work has been, and is being, done by inductive methods, and there have been evolved several laws which claim to be in the nature of physical laws. The early eighteenth-century law of Gregory King relating the supply and the price of wheat, the law of population of Malthus, Engel's law relating to the distribution of expenditure, and Pareto's law relating to inequalities of income are examples of this kind. But none of them possesses that unquestioned authority which the physical laws command.

Abstraction. Economics must, therefore, have re-

course to a much more elaborate method. If it cannot experiment, it can do something analogous to it by intellectual method, that is, by the use of *abstraction* and *isolation*. By these means we can, in our minds, construct 'ideal' conditions. We can abstract from the facts of the real world and we can proceed to study one, two, or more facts in isolation. Clearly, the results which this method will yield will be as 'ideal' as the physical laws elaborated under laboratory conditions. They will only be true with the constantly used qualification of *caeteris paribus*; but provided that the facts for isolated study have been well chosen, and provided that the path back to the complexity of the real world is kept open, there is no reason why these 'ideal laws' of economics should not be useful guides. They will express certain tendencies which can be expected to operate under the influence of those factors which have been taken into account. But they will be modified in reality by all those other factors which the original analysis has neglected.

Different levels of abstractions are possible; we can, roughly, distinguish three such levels in economic inquiry. On the first, or highest level, we start with some assumption which is itself derived from experience, but which is so self-evident that it can be taken as axiomatic. This procedure is analogous to that of Euclidean geometry. The very definition of economics given in the preceding pages represents such an assumption. The subsequent process is to

develop the implications of the original assumption by a process of logical deduction. If the assumption is sufficiently axiomatic, and we have seen that the definition of economics is of that nature, we shall, on this level of argument, arrive at conclusions of universal validity. As we shall see, however, when we apply this method, these conclusions will often be little more than tautologies. The relationships which they establish are general; and they tell us little, therefore, about concrete situations.

On a somewhat lower level of abstraction in pure economic theory, we add to the conclusions derived by the previous method certain hypotheses which we regard as well founded. They represent, in the words of John Stuart Mill, 'real facts with some of their circumstances exaggerated or omitted'. Most of those general data of our society described above belong to this category. Through adding these we can elaborate theories and laws, particularly in the sphere of exchange, which are calculable and precise. But it must be obvious that, on this level, we must exercise a constant watchfulness to see that our conclusions do not become invalidated by a change in the degree of realism of our sociological hypotheses.

Finally, on the lowest level of abstraction, we are dealing with the individual instance. This is the real field of historical study (including, of course, contemporary history), and there is little room for it in a general introductory work. Here, all the circumstances of the case will have to be taken into account,

and inductive investigation will have to be added to the deductions of pure theory.

We see, thus, that in economics, too, there is an interpenetration of deduction and induction. The original assumptions and hypotheses of theory are themselves inductive; and induction should be constantly used in order to check deduction and in order to suggest new problems for theory to tackle.

To sum up, the scientific apparatus of economics does produce certain conclusions of universal validity, but these are so devoid of concrete content that they are not directly applicable to practical problems. The more exact laws of economics, on the other hand, are generally dependent on the inclusion of certain hypotheses, the relevance of which is subject to change.

The concrete situations upon which the economist is often asked to pronounce are generally so complex that he will not only have to employ his scientific apparatus, but will also have to draw on his general knowledge of man and affairs which, though not scientifically precise, will help to give his statements relevance, his judgements wisdom, and his forecasts probability.

PART TWO

THE WORKING OF THE ECONOMIC PRINCIPLE

I. THE THEORY OF CONSUMPTION

I. A DIVISION OF ECONOMIC THEORY

Production, distribution, exchange, and consumption. We will follow the method set out above and begin with the most general propositions. Our starting-point is the definition of economics: a study of the problems that arise in the administration of limited means that have alternative uses. It is always wise to divide the problems one is about to study into certain broad classes; this has been the procedure from a very early stage in the systematic study of economics. The broad divisions first introduced over a hundred years ago, and still much in use to-day, are those of production, distribution, exchange, and consumption.

At first sight this appears to be a very logical division. Under the heading of production, we would study how the means of satisfaction are provided; under that of distribution, we would be concerned with the manner in which the total product is shared among the different groups of individuals who have helped to produce it; exchange would be that part of theory which studies the way in which these individuals exchange with each other the goods and

services which they possess; finally, in consumption, the application of these goods and services to the obtaining of satisfaction would be examined.

Modern economists have decided that this classification is not as useful as it appears. It is, in fact, a classification which was appropriate to the older definition of economics as the study of the production, distribution, and exchange of material goods; for only if we assume that we are dealing with material goods does the division become fairly watertight. Otherwise there is little ground for separating production from exchange, for example, for the transactions with which we are concerned in the latter category also contribute to making available means of satisfaction. Again, the division between production and distribution, and between distribution and exchange, is difficult to uphold. For, in so far as the study of production is concerned with other than technical conditions, viz. with economic ones, it does deal with the quantities of the original resources that are combined in the productive process. This, under the above classification, would be part of distribution, while it could be argued that the purchase of these resources in the market for purposes of production should be part of the study of exchange. It will become clearer as we proceed that this classification is not well adapted to the task of developing the implications of the modern definition of economics.

Consumption and production. A new division, therefore, becomes necessary. According to the

definition we have adopted, the study of economic activity consists of an inquiry into the problems of choice. This can be divided into two parts. The first concerns itself with the way in which given limited means are applied to the satisfaction of wants; this part is called the *theory of consumption*. The second part inquires into the manner in which these means are obtained; and this constitutes the *theory of production*. The significance of this division will become clear as we proceed; but some of its implications may be stated here.

First, it must be understood that we are not distinguishing between different activities. Production covers the study of all activities which make more means of consumption available. The term, therefore, has a much wider meaning than its everyday one which is confined to the making of physical goods, for we also include in it commercial activity (distributive trades, &c.). The production in our sense of, say, a bar of chocolate includes not only the activities of those engaged in making the ingredients and the finished product, but also those who transport them and who deal in them at the different stages from the raw materials to the purchase of the final article in the retailer's shop. While exchange is thus included in the theory of production, it is more convenient, for reasons which will soon be apparent, to study the problems of exchange separately and immediately after, and in conjunction with, the simplest problems of choice treated in the theory of consumption.

The three situations, then, in which we study the

problems of choice are: where there is neither production nor exchange (in the ordinary sense of these words); where there is exchange but no production; and where there are both production and exchange.

In the theory of consumption, we are endeavouring to make generalizations about the manner in which means are used for want-satisfaction. Possession of a stock of such means is assumed. The generalizations which result are applicable, therefore, to the *individual*, whether as a member of society who has (in a manner yet to be studied) obtained these means, or as an isolated being. It is not, at the outset, necessary to assume all the complicated transactions of the industry and commerce of our social system. When we have studied the phenomena of individual consumption we can proceed to enlarge our conclusions so as to include the activities of exchange. In other words, we shall proceed to add the assumption that choice is also effected through transfers between individuals of means of satisfaction. Only then shall we be in a position to study the problems that arise in the process of increasing the number of these means or their availability in time and space. Thus we have consumption followed by exchange followed by production. But it must be understood that the theory of exchange is separated from that of production for purposes of convenience only; the activity which it studies is not to be regarded as necessarily unproductive.

2. WANTS

Wants and needs. We begin with the isolated economy. This term must not be misunderstood. It does not, as is sometimes assumed, refer only to Robinson Crusoe on his desert island. It is also intended to cover the case of a social economy planned by a central authority, or that of the individual in our own unplanned social economy in so far as we only consider his actions in the consumption of a given stock of means. The essence of an isolated economy is, therefore, the existence of a single control over the means, of a single will determining their disposition.

The first elementary fact to notice about the individual (we use that word for brevity) is that he experiences *wants*. They include the most primitive physiological needs as well as all those refined desires to which the more complex life of civilization gives rise. We are not concerned with the origin of these wants or with their 'rightness' or 'wrongness' from any absolute point of view. It is a matter of indifference to us whether these wants are due to snobbishness, extravagance, ignorance, or bad habit. We do not wish to distinguish between the want for bread, for holidays on the Riviera, or for opium. We simply note that human beings experience wants.

But we must further distinguish between a need and a want. We may say that a human being needs, from a medical point of view, a certain diet, or that he needs, from an ethical point of view, certain instruc-

tion and education. But these needs are all determined by certain absolute, objective standards and may not at all correspond to the wants which are experienced by the individual himself. We use the term want to denote neither desire in general nor need externally determined, but an experience of lack of satisfaction which leads to action designed to provide that satisfaction.

Four qualities of wants. Can we discover any general characteristics of wants thus defined? The first characteristic of wants is that they are *unlimited* in number. This statement does not assume that every human being is greedy. But there can be no doubt that, potentially, an infinite number and variety of wants can be added to those which are already clamouring for satisfaction. Even as our capacity to satisfy them grows, we get accustomed to new wants.

Closely allied with the above is the *competitive* nature of wants. This is a characteristic which follows directly from the relative limitation of resources which gave us our definition of economics. Wants compete with one another for the limited resources (with alternative uses) which we possess. They conflict with one another and thus put us under the necessity of choosing which to satisfy. In fact, as Professor Knight has pointed out, we could not conceive of wants in the sense in which we have used the term, unless resources for satisfaction were limited, thus leading to conflict and choice. If any one has difficulty in fully appreciating this point

he has only to think of the many claims upon his time and energy that cannot all be satisfied.

It should be noted that another sense is sometimes applied to the term 'competitive', involving the more directly alternative nature of some wants. Some wants may give way to others; e.g. the want for coffee may give way to that for tea, the want for Devonshire to that for Wiltshire butter. In these cases the want is not specific enough to affect one's conduct with regard to the means of satisfaction. This alternative nature of wants is often important, as we shall see; but for the moment we must confine the use of the term 'competitive' to the general case outlined above.

A third quality of wants is their *complementarity*. Many wants go together: one gives rise to another; the satisfaction of one cannot be completed without the satisfaction of the other. Here, again, we can conceive of complementarity in a narrow and obvious, or in a wide and more remote, sense. Some wants are very closely related. The wants for motor-car and petrol, as those for horse and cart, complement each other. But, beyond this, we must bear in mind the general connexion between all wants which human beings experience. The want for a suburban house may go together with that for a motor-car, with that for certain kinds of clothing, and with the want for a certain quantity of entertainment. It is possible to conceive of a general structure of wants which are related to one another and which mutually condition each other. This complementary character of wants has a bearing

on choice. It implies that choice is never the result of a want in isolation, but of a general system of wants.

As a fourth characteristic we can point to the *recurrence* of wants. The great majority of the wants which we experience will tend to recur even after they have been satisfied. This is obvious in the case of the more primitive wants, e.g. food, sleep. But even more refined wants tend to become habitual, to lead to a recurrent desire for satisfaction. Habits of this kind are constantly being increased. A comparison of the habitual wants of the average twentieth-century European with those of his ancestor of 150 years ago is enough to show the ease with which new wants can be created. And even though some may have been generally discarded, the net total is to-day infinitely greater than it then was. The phrase *standard of living* denotes something of this idea of recurrent wants. It is used to describe a system of wants, the satisfaction of which has become habitual for certain individuals or groups of individuals. From the point of view of the theory of consumption the quality of recurrence in wants is important. It implies that choice will take into account not only wants experienced at any given moment of time, but also the future wants which experience leads one to anticipate. These anticipations, too, will affect conduct; they can therefore, for our purposes, be regarded as wants.

The first law of Gossen. The fifth characteristic of wants is so important that it requires special treatment. When we examine any individual manifestation

of a want, we find that it is *limited in capacity*. It is possible to satisfy it to such an extent that, for the time being, it disappears. Before this happens progressive satisfaction will have led to progressive diminution in the intensity of the want. This fact was first stated in a precise manner about the middle of last century by a German writer, Hermann Heinrich Gossen. He put it in the following form: 'The amount of one and the same satisfaction declines, as we proceed with that satisfaction, until satiety is reached'; and this statement is now known as *the first law of Gossen*. Any one can easily convince himself of the truth of this law of diminishing want-intensity. The usual examples, such as the one of the progressive consumption of, say, slices of bread, can certainly show that, as we begin to satisfy it, the intensity of a want decreases until it becomes zero. There is no want to which this law is not applicable, provided that we bear in mind that we are speaking of individual want-manifestations giving rise to conduct and not of any general kind of want. Thus, it is not the desire for bread in general which we have in mind in the first law of Gossen, but any concrete want for bread at a given time which leads to the consumption of bread in order to satisfy it. This latter want is undoubtedly satiable and the law applies.

But this is not the form in which the first law of Gossen becomes of importance for our purpose. We have to interpret the observed diminution of want-intensity in a different and more appropriate manner.

Since wants compete for satisfaction (because means are limited and have alternative uses), the intensity of any want which is being satisfied declines *in relation* to other wants; that is to say, other wants become more clamorous for our attention as we satisfy any given want. It is in this form that, as we shall see presently, Gossen's law becomes fundamental for the formulation of a theory of consumption.

It remains now to point out a qualification of the validity of this law. From the way in which it was formulated, it might be thought that want-intensity declines by infinitely small stages. In fact, wants are not infinitely divisible; there is a certain minimum intensity, which has been called the *minimum sensible*, below which a want has no effect on conduct. But at this stage in our analysis this fact does not seriously affect the law. Thus, we need not be unduly concerned because certain wants seem to increase in intensity with the first few doses of satisfaction and do not begin to diminish until later. There is another seeming exception—that of the collector whose appetite seems to increase when he first begins to collect. This, on second thought, turns out not to be an exception at all. For Gossen's law assumes a certain 'given' state of tastes. And when the passion for collection has been aroused after one acquisition, the individual's tastes can be said to have altered.

3. GOODS

Classifications of goods. We must now proceed to examine in detail the means of satisfaction, the counterpart of wants. Everything that an individual regards as capable of satisfying a want we call a *good*. All material and even non-material objects (position, power, &c.) and all personal services are goods. We can classify goods in three different ways: the first two referring to their physical nature, the third to their quantitative relationship to wants.

In the first place, there are goods which are directly and immediately capable of satisfying human wants. A loaf of bread, a pound of butter, a suit of clothes, a house, a gramophone, the personal services directly rendered by teachers, doctors, garage mechanics, or domestic servants—all those are goods ready for consumption. We call them *consumption goods* or goods of the *first order*. The possession of goods of this kind will, at least in the opinion of the individual concerned, secure the satisfaction of wants.

There are also goods which only indirectly serve to satisfy our wants. If we take bread as our example of a consumption good, we can put into our second category all those goods which have gone to the making of bread: flour, water, yeast, salt, the oven and the things it is made of, the fuel for heating it, and all the other accessories. These goods we call *production goods*, or goods of a *higher order*, because they have to be combined in the productive process

in order to result in those goods which directly satisfy wants.

The term 'goods of a higher order' is used to convey the picture of a structure of goods, the first of which, when viewed by the individual experiencing a want, is the finished consumption good, the second, the goods which are only one stage removed, and so on. The importance of this picture lies in the fact that it makes us realize that, in order to satisfy wants by means of production goods, we must possess *all* the goods which the process of production requires, we must possess them throughout the process, and we must possess the knowledge and the labour necessary to carry through the process. We shall study the implications of these facts in the theory of production. It should, at the same time, be clear that the distinction between consumption and production goods is not a permanent, unchangeable one. There are many goods, water, for example, which may be used either for direct want-satisfaction or for productive purposes. The distinction remains, nevertheless, important, for, as we shall see, the decision whether to use a particular good in the one way or in the other presents a considerable problem.

A further classification, affecting both consumption and production goods, is possible: this takes into account the number of times which the good is capable of satisfying a want. There are goods whose capacity of want-satisfaction is exhausted in a single use; goods which can render a series of satisfactions and lose their capacity only gradually; and, finally, goods which are

permanently capable of rendering satisfaction. For want of better terms we can call the first *perishable*, the second *durable*, and the third *imperishable* goods. But it must be remembered that these words have somewhat different meanings in ordinary language. Examples of these three types will readily come to mind. All food-stuffs belong to the first category. Clothes, on the other hand, will yield a series of satisfactions over a period of time, though they will wear out eventually. Finally, a country estate may go on yielding satisfaction to successive generations without losing the capacity to do so in the future.

Free goods and economic goods. Our two classifications have referred to the technical and physical characteristics of goods. There is, however, yet another characteristic: their quantitative relationship to wants. With this we have already become acquainted since it helped us to arrive at a definition of the subject-matter of economics. There are goods which are not limited in relation to wants; these we call *free* goods. Light, air, water, timber, land may, in certain conditions, be free, for they do not necessitate choice and economy.

Goods may be limited in relation to wants: they may be scarce; then they are called *economic* goods, for their disposition becomes a problem of choice. It is only with these goods that we are concerned in economics. But, again, it must not be forgotten that individual goods may easily pass from one class to another. Air in crowded rooms, water in cities, timber and land

in densely populated areas, from being free have become economic goods. Changes in habit may also make economic goods into free ones; conversion to vegetarianism, legislation against the private motor-car, a campaign against dress shirts, may quickly alter the quantitative relationship between goods and the wants for them.

All economic goods constitute *wealth*. Wealth, thus, possesses the following characteristics: it is capable of satisfying wants, either directly or indirectly, either in a single use or in a series of uses, and it possesses scarcity. It includes all material as well as all non-material goods and services, but it excludes those which are free. Elaborate arguments over more precise definitions of wealth are generally useless, though further distinctions may sometimes be necessary, such as between the wealth of an individual and collective wealth. The latter is not part of individual property, but individual satisfaction is derived from it. Roads, parks, an efficient legal and educational system, and all the other social institutions which are collectively owned, would be included under this heading.

4. UTILITY

We must now proceed to study in greater detail the want-satisfying capacity which we have found to be the quality of a good. At the beginning of the last section it was made clear that this quality depended on the view which the individual took. If, therefore, we now introduce the term *utility* to denote the

capacity of want-satisfaction, it must be understood that this term is not used in its everyday sense. It does not refer to an objective quality, nor to any ethical 'rightness' or 'usefulness', but expresses a relationship between a good and an individual. It is so important to realize the true nature of this concept as used in contemporary economic theory that it is worth while treating it at some length.

We know that there are certain objective qualities inherent in external things which enable them to satisfy human wants. A loaf of bread contains certain chemical elements combined in a certain way which make it able to appease hunger. A table, because of the kind of material of which it is made, and because of its dimensions, can serve as a resting-place for books or food. Because of its design or colour it may serve to satisfy other, more refined, wants of an aesthetic character.

When we speak of the utility of an object in ordinary language we generally have in mind the kind of quality just mentioned. Sometimes we make the word refer even more definitely to an objective standard; we then contrast 'utilities' with 'luxuries'. Needless to say, not everybody will have the same idea as to what is useful and what is useless, but the fact that at any given time certain criteria of utility are widely held leads us in everyday life to use this word to describe an objective quality.

After our discussion of the characteristics of wants it should be clear that modern economics tries to avoid

any such reference to objective rightness or wrongness. In our use of the term, utility must, therefore, be shorn of any ethical or other objective implication. It must become purely subjective and relative. Thus when we say that a good possesses utility, we simply mean that it is regarded by some individual as capable of satisfying a want which he experiences. Utility expresses a relationship between a subject and an object from the point of view of the problem of want-satisfaction.

Thus understood, the concept should leave no room for the paradoxes which are sometimes so puzzling. It should now be clear why, economically speaking, meat and beer possess utility even though there exist vegetarians and total abstainers. And it should no longer cause surprise when the economist says that opium possesses utility (apart altogether from its medicinal uses) even though addiction to drugs is almost universally regarded as harmful and 'wrong'.

It is unfortunate that we should have, once again, to employ a term in a sense so utterly different from that of everyday usage. Some economists have therefore put forward other terms.

For example, it was thought that *desirability* would better express the essentially relative nature of the concept. But it is difficult to see why desirability should be superior in this respect to utility. There are many commonly accepted views as to what is desirable just as there are as to what is useful, and these are always liable to intrude on the use of either term in economics.

A better term is *significance*. It stresses both the relationship between subject and object and makes it possible to include a reference to individual conduct. We could, thus, say that an economic good is everything which possesses significance for individual choice. The word has not, however, been able to oust the commonly accepted term, utility, and we shall therefore keep the latter. It is essential to bear in mind its peculiar meaning; it is subjective, it is relative, and it ignores the inherent qualities of goods (though, admittedly, it is on them that ability to satisfy wants depends).

Diminishing utility. In a sense, therefore, utility is nothing more than a duplication of the concept of want-intensity which we have already met; and the statements made about the latter will also apply. In particular, the first law of Gossen can be formulated with regard to utility, and is then given the name of the law of *diminishing utility*. It is of the utmost importance to insist on the relative nature of diminishing utility; and reference to it has already been made in connexion with wants. What the principle of diminishing utility really expresses is a relationship between the quantity of a good and its utility. If we suppose an individual to possess a uniform supply of a good, which is capable of alternative uses, the quantity which he possesses will determine the uses to which he will be able to put it. He will, in other words, apply successive units of the good to wants of diminishing intensity, and the size of his stock will determine

the least intensive want which he will be able to satisfy. This statement remains valid whether we are thinking of the same generic want or of wants that appear to be of different kinds, though satisfiable by the same good.

In the well-known example of the boy consuming successive slices of bread and butter we must, in order to see clearly the nature of Gossen's law, conceive of each successive want for a slice as a separate want-manifestation, each less intensive than the preceding one. The number of slices on the plate will determine the least intensive want which can be satisfied or the least intensive use to which the supply can be put.

The other classical example, of the isolated individual who possesses a uniform supply of bushels of corn, makes this even clearer. With one bushel he can probably satisfy nothing but the most intensive want: that for the food necessary to maintain himself for a period of time. The possession of a second bushel may allow him more abundant nourishment or the same amount of food spread over a longer period. A third bushel may be used as seed in order to provide for a more distant future; a fourth may serve as feeding stuff for poultry and thus contribute to a more varied diet; and a fifth may enable the comparatively feeble want for a distilled spirit to be satisfied. It is easy to see that the intensity of these wants declines in the same manner as the successive wants for bread and butter. The utility of the good, therefore, since it is no more than a duplication of want-intensity, also declines.

Marginal utility. But it would be very misleading

to conclude that the utility of any bushel of corn is different from that of any other. We have assumed that all the portions of the supply are uniform, and that they are capable of alternative uses. As the most intensive want is satisfied, it diminishes in relation to the next intensive want, and so on. When we reformulate this by saying that the utility of a good diminishes, this must not be construed as meaning that in a given supply each portion has a different utility. If only one unit were available, the utility would be of one intensity; if two units, the utility of each unit would be smaller, since a less intensive want could be satisfied out of the available supply; and so on. We are in this way led to the concept of a *margin*. The least intensive want which a given supply of a good permits one to satisfy is the *marginal want*; the portion of the supply which has to be applied to the satisfaction of the want is the *marginal unit*; and its utility is the *marginal utility*.

The concept of the margin is of great importance. A little reflection shows that it flows directly from the relative limitation of resources which leads to economic activity. When we say that a want is marginal, we wish to make more precise the dividing line between the wants which can be satisfied and the wants which, because of the scarcity of goods, have to remain unsatisfied. An enlargement of the stock of goods beyond the marginal unit will provide satisfaction of wants which previously were extra-marginal; and since these wants are less intense in relation to the preceding ones the marginal utility of the good will diminish.

A further conclusion follows. The utility of any unit of a stock of goods can never be greater (or less) than the marginal utility. In a uniform stock all units are interchangeable. Even if we assume that the individual bushels of corn in our previous example had originally been earmarked for specific uses (of different urgencies), the loss of one bushel will clearly mean the loss of the least urgent use. If the first bushel is lost, it is inconceivable that the individual should forgo the satisfaction of the most urgent want for which the bushel was destined, viz. that for food. At once, by a process of substitution, all the other bushels will move up the scale and the want which will have to remain unsatisfied will be the least urgent one, viz. that for spirit. Thus, the utilities of all the portions of a uniform stock are equal and are determined by the marginal use to which the stock can be put, i.e. its marginal utility.

5. MAXIMUM OF SATISFACTION

The structure of wants and the structure of goods.

We have so far developed a theory of choice for the isolated individual who disposes of a uniform stock of goods with alternative uses. We have seen that he will apply this stock to the satisfaction of his wants according to the order of their intensity, beginning with the most and descending to the least urgent. The utility of the good will then be dependent on its quantity and on the individual's wants.

We must now proceed to examine the more realistic

assumption that the individual disposes of different quantities of a *variety* of goods, which cannot be substituted for one another. In the first place, the previous conclusions will still hold good, viz. each individual utility will still be dependent on the individual want and on the quantity of the individual good.

But these will no longer be the only determinants. We have already seen that wants are complementary; many wants call forth others, and some wants can only be satisfied in conjunction with others. If want-intensities are interdependent, utilities must be so too. The utility of a good appears, therefore, to be dependent not on the quantity of that good alone, but also on the quantities of all the other goods which the individual possesses in relation to all the wants which he experiences. Just as it was necessary to think in terms of a structure of wants, so it is necessary to envisage a structure of goods with which it harmonizes. The general interdependence of wants is, thus, accompanied by a general complementarity of goods.

The second law of Gossen. The individual will have to arrange his choice in accordance with this complementarity. To this end, he will have to aim at a more or less uniform satisfaction of wants. The way in which this is achieved was stated by Gossen, in his *second law*, as follows: 'If it is impossible to gratify all wants to the point of satiety, it is necessary, in order to obtain maximum satisfaction, to discontinue the satisfaction of different wants at the point at which their intensity has become equal.' It would be un-

economic to satisfy one want to the point of satiety and to discontinue the satisfaction of another while its intensity was still great. The highest satisfaction is reached by an even level of utilities, i.e. by an equalization of marginal utilities.

While this law follows from the first law of Gossen and from the complementarity of wants, it requires some qualification. Its realization in practice is impeded by the already noted fact that wants are not always infinitely divisible. In the same way, it can be said that goods are not generally infinitely divisible either (this applies particularly to many durable goods; motor-cars, houses, &c.); and for that reason the second law of Gossen must be understood as the statement of a tendency only.

Consumption in the isolated economy. We are now in a position to summarize our conclusions concerning the conduct of an isolated economy and, in so doing, to elaborate a little further a number of points. Economic conduct consists of choice necessitated by the existence of wants which compete for limited resources. These wants will, through the acts of choice, be arranged in a certain scale of intensity, and satisfaction will proceed according to that scale. Maximization of satisfaction demands that as uniform as possible a level of marginal utility should be achieved, since with progressive satisfaction the intensities of wants decline in relation to other wants.

The competitive nature of wants is seen most clearly when we assume that the means of satisfaction consist

solely of effort. As we expend it in order to satisfy a want, we shall experience two sets of sensations which it was customary to describe as pleasure and pain. Effort, it was said, produced utility, but it also involved disutility. Whatever may be the psychological truth about the difference in these sensations, it should be clear now that, economically, they represent the same type of phenomenon. If, for example, we use effort to gather fruit, there will be a relative decline in the utility of fruit and a relative increase in the utility of all the other alternative uses of effort, including leisure. Once the relative nature of utility is understood, it is not necessary to maintain in use either the pleasure-pain comparison of some earlier economists or the concept of disutility which is still met with to-day.

A problem closely allied to the above is that of the measurability of utility. To form the basis of choice, utility must be conceived of as a magnitude. But can it be said that utility is measurable? If we were to ask our owner of the five bushels of corn what was the utility to him of corn, he would be unable to give us an answer. In so far as we are concerned with the isolated economy, it must be clear that utility, being the same as want-intensity, is an intensive and not an extensive magnitude of a psychical character and that it cannot, therefore, be measured.

But is this necessary? All that the individual requires for the economic disposal of his resources is a conception of greater and less want-intensity (and

utility). He merely needs to *compare*, to establish a scale of preferences, in order to be able to evaluate the means of satisfaction. And we have seen that this evaluation will be dependent upon the marginal use to which a stock of goods can be put.

It may be absurd to ask our individual about the value to him of corn, since (in the absence of other food) we should be asking him to assess the value he puts on his own life. It is not absurd to ask him about the value of one bushel of corn, since he will then be asked to *compare* the intensities of two wants; and given the total stock, it will be the marginal use which will determine the value to him of any unit. He can say that if he had three bushels of corn he would value each one more than if he had four, and so on.

With the aid of the marginal utility concept, some well-known paradoxes can easily be disposed of. As will become even clearer in the course of our analysis of exchange, it is no longer surprising that to some of the most useful goods, such as water and air, we attach little value. They are either free goods which do not involve any economy at all, or they are present in such quantities in relation to wants that a single unit possesses little significance for conduct. It is clear that water is indispensable to human life; but to obtain an additional amount of it involves a very small sacrifice of alternative opportunities of satisfaction.

A link with exchange. By stressing the alternative uses of our limited means of satisfaction and the tendency expressed in the second law of Gossen we

have also established a link with the theory of exchange. For, in the first place, the economic activity of the isolated economy involves a kind of internal exchange—the exchange of the satisfaction chosen, for the alternatives foregone. The maximization of satisfaction in Gossen's law clearly involves considerable exchange of that kind.

The second law of Gossen points to yet another possibility. To maximize satisfaction means to harmonize a structure of goods with a system of wants so that a point may be reached at which want intensities are equal. Exchange of goods between individuals is a powerful aid to this end. When some goods are present in such quantity that some wants can be satisfied to a very low degree of intensity, while other goods are so scarce that the wants to which they can be applied must continue to remain very urgent, the individual will endeavour to exchange the abundance of the one good in order to diminish the scarcity of the other. To the problems which this involves we must now turn.

II. THE THEORY OF EXCHANGE

I. THE CONDITIONS OF EXCHANGE

IT may appear strange that we should only now be reaching what has generally been regarded as the central problem of economics: the explanation of quantities and ratios of goods exchanged.

The interest in this problem is due to the importance which the process of exchange has acquired. In an earlier chapter we have seen how certain economic and institutional changes have resulted in an ever-widening market in which all economic goods, including the original resources of mankind, were being habitually exchanged. It was inevitable, therefore, that the phenomena of the market should have come to occupy the central place in economics. In recent times, however, the desire to make economics more scientific, that is, to make its conclusions more widely applicable, has led to the search for simpler assumptions, free from the limitations of any particular set of social data. In the preceding sections we have met the results of this search; in the theory of consumption the assumptions are the simplest one could find: an isolated economy and a given stock of goods.

We have seen that conduct under these assumptions can already be regarded as a kind of exchange, and we have noticed that the maximization of satisfaction may be made impossible by the composition of the stock of goods possessed by an individual. The next

step natural to our analysis is a study of the conduct of individuals who are endeavouring to change that composition through exchange with others who are similarly placed.

Earlier, we saw how exchange arises from division of labour. But we see now that, theoretically, it is not necessary to assume division of labour in order that exchange may become possible. Indeed, at this stage of our inquiry, we need not be concerned with differences in people's ability to produce goods. All we need assume is the possibility of increasing want-satisfaction by parting with certain goods in one's possession in exchange for others in the possession of somebody else.

Differences in marginal utility ratios. It remains now to inquire in what conditions this possibility will exist. Some indications have already been given. An individual may possess a stock of goods which enables him to satisfy one want to a considerable extent while another want has to remain almost unsatisfied. For the sake of simplicity we will assume that he possesses two goods with, to him, unequal marginal utilities. One will occupy a higher, the other a lower, place on his scale of preferences. That being so, he will be prepared to part with portions of the latter in exchange for portions of the former until their marginal utilities are equal.

Suppose now that another individual also possesses these two goods but their respective places on his scale of preferences are reversed. He, too, would be prepared to exchange; and clearly a meeting of these two

individuals would make exchange possible. We see, therefore, that the first conditions for exchange between two people are (*a*) that they should both possess the goods to be exchanged, and (*b*) that these goods should occupy different places on the scale of preferences of each individual.

Before going farther it is important to make one point clear. Utility is assumed to represent in some way an individual's estimate of the satisfaction which he derives or expects to derive from a good (for desire and satisfaction need not correspond). Such estimates are entirely personal and intensive; and it does not follow at all that the satisfactions derived by the participants in exchange are equal. The satisfaction which I derive from a concert can only be evaluated by me; the satisfaction derived by the organizers from the money I pay can only be evaluated by them. There is no objective standard which would enable us to equate the two.

Wicksteed expressed this by saying that there was 'no bridge' between the subjective evaluations of different individuals. It is doubtful whether that particular phrase is quite appropriate. For many readers may say, even if we admit that utility cannot be measured, and that, therefore, the utilities of the same good to different individuals cannot be compared, there is yet the undeniable fact that individuals do exchange in certain quantitative relations and that they do thus appear to find a means of expressing their scales of preferences. In the act of exchange a

'quantification' of utilities takes place. The act of exchange does, therefore, provide a bridge for two individuals' subjective evaluations. How is this 'quantification' achieved?

The marginal rate of substitution. In the isolated economy, the problem of 'quantification' does not appear to arise in a very acute form. But even here conduct is essentially in the nature of exchange (the substitution of one alternative for another), and it possesses, therefore, a quantitative character. The substitutional aspect of choice becomes more obvious when exchange proper is considered. But once we stress this substitutional character of exchange it becomes clear that all we need to know in order to explain an individual's conduct, is his *marginal rate of substitution* between any two goods, given his scale of preferences and the amounts of goods he possesses.

This new term, which has been introduced recently by J. R. Hicks and R. G. D. Allen, has the great advantage of being definitely quantitative. By adopting it, it seems even possible to do without utility altogether; and at the end of this book some indication will be given of the arguments put forward for and against the retention of this term. Since, however, we have, in this text, kept to the usual analysis which employs the concept of utility, it is necessary to point out the relation between the new term and the old.

We have found that the utilities which goods possess for individuals are not measurable. We have also

is a heavy smoker, would only be prepared to part with 1 oz. of tobacco if he could get at least 1 bushel of corn for it. B, on the other hand, because he has very little corn and/or because smoking is to him only a pleasant luxury, would be prepared to part with, at the most, $\frac{1}{2}$ bushel of corn in order to get 1 oz. of tobacco. In these circumstances exchange is impossible; the ratios of their marginal utilities is such that exchange could not benefit both parties, and no bargaining can take place. If B agreed to A's terms he would be worse off than he was before, since he would be exchanging a good with a higher for one with a lower marginal utility, and vice versa for A.

Suppose, however, that A would be prepared, in order to get 1 bushel of corn, to part with 2 oz. of tobacco, while B would be prepared to give 1 bushel of corn if he could get at least $1\frac{1}{2}$ oz. of tobacco. Exchange is now possible. The *ratio of exchange* must clearly lie between $1\frac{1}{2}$ oz. to 2 oz. of tobacco for 1 bushel of corn. But can we say anything more than that? Can we say more precisely what the ratio will be? At either of these ratios an equilibrium is possible. Each will feel satisfied at a ratio of $1\frac{1}{2} : 1$, $2 : 1$ and, therefore, at any intervening ratio. At any ratio within the two extremes, the condition that the relation between the marginal utilities of the two commodities must be equal for both participants will hold good.

We find, therefore, that the problem of isolated exchange is *indeterminate*, i.e. that there are many possible ratios (within certain limits) at which ex-

change may take place. The determination of the ratio in practice will depend on many circumstances, the most important of which are the relative astuteness, perseverance, and bargaining power of the two parties. If A is superior in these respects, the ratio will lie nearer to $1\frac{1}{2} : 1$; if B is more successful, it will lie nearer to $2 : 1$.

Bilateral monopoly. The theoretical importance of isolated exchange lies in the fact that it represents the first case to be met when we pass on from the consideration of the isolated individual. In practice it may be said that, as depicted here, isolated exchange is of little importance. 'Swopping' among schoolboys is its most frequent realization, but much more complex exchange relations are typical of modern economic life. There is, however, one sense in which isolated exchange can be said to occur and to be of importance in practice. That is where powerfully organized groups face each other as sole seller and sole buyer. This situation, known as *bilateral monopoly*, may be found in practice in the case of wage bargaining between associations of employers and employed. Without going into the details of the problem at this stage, it can be said to be akin to that of isolated exchange. The ratio of exchange will therefore be indeterminate over a certain range; and the fixing of it, in any given instance, will depend on the relative bargaining strength of the two sides.

Unilateral competition. The case of more than two individuals wishing to exchange two commodities

is little different. We will assume that two individuals, A and A_1 , wish to dispose of tobacco in exchange for corn, and one individual, B, is prepared to accept tobacco for corn. The willingness to exchange of A is $1\frac{1}{2}$ oz. of tobacco for 1 bushel of corn; that of A_1 2 oz. for 1 bushel. B is prepared to part with 1 bushel for anything over 1 oz. We can see at once that exchange will be possible and that the ratio of exchange will lie somewhere between 1 oz. and 2 oz. for 1 bushel. Can we determine the ratio more precisely? Only to the extent that we can say that B is bound to discover to what ratio A is prepared to go. The lower limit will at once move to $1\frac{1}{2} : 1$; but since A_1 will outbid A, the latter will be excluded from the bargain altogether. It will then depend on the relative skill of A_1 and B whether the ratio is nearer to $1\frac{1}{2} : 1$ or $2 : 1$. In this case, while the actual ratio cannot be determined in advance, we can, at any rate, fix the limits a little more precisely.

The two examples we have so far examined do not carry us very far. Their limitation is the assumption that the quantity of only one of the two goods is variable (in both cases that of tobacco), while the quantity of the other (corn) which is to change hands is already determined. This is an unreal assumption, because we know that in actual life the quantities which different people are willing to exchange will vary, and they will vary for the same individual from time to time according to his general system of wants.

More precisely, we know that if A is, in general,

willing to exchange X for Y, it is so because they occupy different places on his scale of preferences. And they do so, other things being equal, because of the relative quantities of each in his possession. If he actually embarks upon exchange, the relative quantities are altered and so are their relative marginal utilities; and exchange will, as has been pointed out, proceed until the marginal utilities are proportionate to the ratio of exchange. But the exact quantity of X which A is willing to give up for Y will depend on the ratio of exchange. At one ratio he may be willing to part with much, at another with little. In other words, we must not continue to assume fixed quantities, but must discover what the ratios will be, since we now know that the ratios will determine the quantities.

Value, demand, and supply. Before we proceed farther it is advisable to clarify the problem by the introduction of several new terms. In the first place it is necessary to point out that the term *value* is often used instead of ratio of exchange. It was customary to call this exchange-value and thus to distinguish it from value-in-use—the subjective evaluation which we have already met. It is customary, nowadays, to refer to it simply as value. We call the value of a good (corn) the amount of any other good (tobacco) for which a unit of it is exchanged. We shall see, however, that the term value will have to give place to another term when the more complex cases of the exchange of more than two goods with the aid of money are considered.

It is useful also to begin to speak of an individual's

demand for one good and his *supply* of another in exchange for it. An exchange consists, therefore, for each participant of both a demand and a supply; and without yet embarking upon a study of the monetary aspect, we can describe them as purchase and sale. A's demand for X is, then, his willingness to buy a certain quantity of it at a certain ratio (i.e. through the supply of a certain quantity of Y); and B's supply of X is his willingness to sell a certain quantity of it at a certain ratio (i.e. through the demand for a certain quantity of Y). It is purely a matter of choice, at this stage, whether we call A the buyer of X and B the seller of X, or B the buyer of Y and A the seller of Y. Later we shall find that the intervention of a medium of exchange makes it convenient to choose one particular terminology.

We can now repeat summarily, and in our new terminology, our previous two cases. We first had one buyer and one seller, a certain quantity to be exchanged, and two ratios relating respectively the demand for, and the supply of, that quantity. A, whom we choose to call the buyer, demands 1 bushel of corn for which he would be willing to pay 2 oz. of tobacco. B, whom we choose to call the seller, is willing to supply 1 bushel of corn for $1\frac{1}{2}$ oz. of tobacco. The actual ratio will be between these two. As A gets more corn (and parts with more tobacco) the marginal utility to him of corn in relation to that of tobacco will decline. The opposite is true of B. The two limits we have arbitrarily assumed represent the points at which

this relative diminution of marginal utilities (corn for A and tobacco for B) make further substitution undesirable.

If we assume the quantities to be exchanged to be variable, there is nothing in our data to enable us to say what the total quantity exchanged will be. We do know, however, that, whatever the total, when exchange ceases the ratio of the marginal utilities of the two goods must be the same for both participants. If we imagine exchange to proceed, step by step, with fractional quantities, the ratios may vary for each partial exchange, because of the changing ratios of marginal utilities to buyer and seller. The average of all the ratios of exchange—which will represent the average ratio for the total quantity bought and sold—is, as we have seen, indeterminate, though we have also seen that it must lie between two definite limits, which depend upon the scales of preferences of A and B.

Discrimination and uniform ratio. Little is altered in our second case, that of a number of buyers and one seller, or vice versa, except that the effect of variations in the quantities exchanged can be more clearly brought out. Where the seller has only one unit of a good, or one indivisible good, to offer, the buyer who (on account of his marginal utility ratios) is prepared to make the highest bid will be the only one to bargain with the seller. Here the situation is analogous to that of our previous case of isolated exchange.

Where, however, the seller is offering several units of a divisible good, the determination of both quantities and ratios is slightly more complicated. There are two possibilities: the seller is offering successive units to be bid for by the buyers, or the seller is offering the whole supply at once. Suppose, for example, that he is offering, one by one, five units of a good, and that he is met by a number of potential buyers, whose demand is for one unit each at varying ratios. We can imagine these buyers arranged in an order of descending ratios, the 'strongest' bidder coming first, the 'weakest' last. Thus we might have A_1 willing to buy 1 bushel of corn at 5 oz. of tobacco, A_2 at 4, A_3 at 3, A_4 at 2, A_5 at 1, A_6 at $\frac{1}{2}$, A_7 at $\frac{1}{4}$, and so on. The evaluations of B, the seller of the 5 bushels of corn, expresses itself in a willingness to sell each bushel for 1 oz. of tobacco. We suppose him to be inviting tenders for each successive bushel. For the first, A_1 , anxious to secure his supply, will offer 5 oz.; he will succeed and leave the field to the others. Then A_2 will obtain his supply for 4 oz., A_3 for 3 oz., and so on until, after A_5 has obtained his for 1 oz., the supply is exhausted and the remaining buyers are excluded. The ratios of exchange lie between the evaluation of the seller, as the lowest, and the evaluation of the strongest buyer as the highest limit. If there had only been 4 bushels available and the evaluation of the seller remained at 1 : 1, the lower limit would lie at 2 : 1, the ratio offered by A_4 , who is the last buyer to be included in the deal. We can call him the marginal buyer.

If the whole supply is offered at once and all buyers have a chance to bid for it, the result will be different. We will assume that B, the seller, wants at least 1 oz. for 1 bushel, and offers 4 bushels. The ratios which the 7 buyers are willing to offer for 1 bushel each are 5, 4, 3, 2, 1, $\frac{1}{2}$, and $\frac{1}{4}$. The demand of 4 only out of the total of 7 buyers can be satisfied, though at least one out of the 3 who must be excluded— A_5 —will be able to bid in the first place since his offer is as high as the lowest which the seller is willing to accept. But he will at once be outbid by the others. Once the ratio has risen above 1 there is no further inducement to any of the buyers to offer more, since they can all be satisfied out of the given supply. There must, therefore, emerge a uniform ratio at which all buyers who are admitted to the deal will buy. That ratio in our example must be higher than 1, but it cannot be higher than 2, for then another buyer (A_4) would be excluded and the seller could not dispose of his stock. The ratio, then, will be limited above by the offer of the marginal, and below by that of the first extra-marginal buyer.

It is not difficult to make the changes which altered assumptions necessitate. If, for example, instead of 4 bushels 7 had been available, then one more buyer would have obtained 1 bushel, and the ratio would have fallen to a lower level. But so long as the seller's minimum ratio remained at 1 two buyers would have remained excluded and the total supply would not have been cleared. Again, it is not necessary to assume

that the seller's supply is fixed beforehand. It, too, may vary according to the ratio he can obtain.

The case we have just been considering, that of a single seller and a multitude of buyers, is of considerable practical importance, and is known as *monopoly*. We are already in a position to indicate at least one point in the policy of a monopolist. We have seen that he may offer his entire supply and allow the actions of the buyers to establish a uniform ratio at which all the exchanges will take place. But we have also seen that he might make offers of individual portions of his supply, and he might then succeed in making separate bargains with each buyer. Where buyers' demands are not all marginal, i.e. where buyers make different offers, the second method would give the monopolist a larger aggregate than the first (this is so in the example given above). In such circumstances, the monopolist will endeavour to split up his buyers in this way and practise what is known as *discrimination*. He will clearly be successful only in those cases in which it is not possible for his buyers to re-sell to one another. The details of other aspects of a monopolist's policy we must leave for a later stage.

3. BILATERAL COMPETITION

The analogy of an auction. We will now assume that there is more than one participant on both sides, though we still maintain the assumption of two commodities. For purposes of convenience, we will call them buyers and sellers. We will simply enlarge our

previous example and suppose that the buyers, $A_1, A_2, \dots A_7$, wishing to buy 1 bushel of corn for 5, 4, $\dots \frac{1}{4}$ oz. of tobacco respectively, are faced by 7 sellers, $B_1, B_2, \dots B_7$, anxious to sell 1 bushel of corn each for $\frac{1}{4}, \frac{1}{2}, 1, 2, 3, 4, 5$ oz. of tobacco respectively. These ratios represent, as we know, the maximum (or minimum) at which, given their scales of preferences, each buyer (or seller) is willing to effect exchanges. It means, of course, as we noticed in our previous examples, that exchanges will not necessarily take place at these ratios; both buyers and sellers will be anxious to take advantage of lower or higher ones. In our example, then, we have the following table of buyers and sellers arranged in accordance with their 'strength'.

Buyers	A_1	A_2	A_3	A_4	A_5	A_6	A_7	
	5	4	3	2	1	$\frac{1}{2}$	$\frac{1}{4}$	} oz. of tobacco for 1 bushel of corn.
	$\frac{1}{4}$	$\frac{1}{2}$	1	2	3	4	5	
Sellers	B_1	B_2	B_3	B_4	B_5	B_6	B_7	

Two questions have to be answered: how many bushels of corn will be exchanged or, in other words, how many buyers and sellers will participate in the deal; and what will be the ratio. In our present example we are no longer concerned with the possibility of successive offers. It is of the essence of the case that both total demand and total supply appear simultaneously. Let us assume that an auctioneer intervenes and 'calls out' a ratio. He begins with the low price of $\frac{1}{4}$; 7 buyers will at once offer to conclude a deal. But at that ratio only 1 bushel of corn will be forthcoming. This discrepancy will at once induce 6 buyers

whose maximum offers exceed $\frac{1}{4}$ to bid more. The auctioneer calls $\frac{1}{2}$. At this ratio 6 buyers wish to buy, but only 2 sellers are forthcoming. The bid goes up. At 1, 5 willing buyers are faced by only 3 willing sellers. Can the ratio alter after that? If it rises to 3, 5 sellers will meet 3 buyers. The ratio must, therefore, lie below 3. At a ratio of 1, as we have already seen, the buyers will exceed the sellers. We conclude that the ratio must lie between 1 and 3. It will, in fact, be 2, and 4 bushels of corn will be exchanged. At any other ratio we should get a discrepancy between buyers and sellers which would mean either that some buyers, still able to buy at the ruling ratio, would be unable to buy, or that some sellers would be unable to sell. The uniform ratio we have found is thus the only one at which buyers and sellers can buy and sell consistently with their own evaluations. It is, moreover, a ratio at which a certain equilibrium rules, since any disturbance, upwards or downwards, must bring about its own correction: the ratio will return to its previous level.

The marginal pairs. This ratio is established by the actions of buyers and sellers endeavouring to obtain the best possible bargain up to what we might call their final 'reserve' ratios. This means that some buyers and some sellers will be excluded. But the presence of those whose offers were near the eventual ratio will have influenced its formation. The ultimate margins between which the ratio must lie are given by that offered by the marginal buyer (A_4) and that

demand by the first extra-marginal seller (B_5), and that offered by the first extra-marginal buyer (A_5) and that demanded by the marginal seller (B_4). Their offers are of importance, and those of all the others who are above or below do not count. We can see this more clearly if we assume a wider spread of individual demands and offers. If we take, for example, the reserve ratios of A_1 , A_2 , A_6 , and A_7 as 500, 50, $\frac{1}{100}$ and $\frac{1}{1000}$ respectively, and those of B_1 , B_2 , B_6 , and B_7 as $\frac{1}{1000}$, $\frac{1}{100}$, 50, and 500, this will not alter the ratio, which must still lie between the evaluations of the marginal pairs.

Consumer's surplus. These marginal pairs are alone of importance in determining the ratio of exchange. Since, however, the ultimate ratio is uniform for all buyers and sellers, it would appear that each one derives a different advantage from the exchange. If A_1 , that is to say, was prepared to go to 5 or even 500 oz. of tobacco for 1 bushel of corn, and he obtains it for 2 oz., he derives a greater advantage from the exchange than A_2 , who would have been willing to pay 4 oz. only, and so on. Similarly B_1 , who would have been prepared to sell for $\frac{1}{4}$ or even $\frac{1}{1000}$ oz., gets 2 oz., while B_2 , who gets the same, would have been prepared to sell for $\frac{1}{2}$ only.

This kind of argument has led to a concept which has been known since Marshall as *consumer's surplus*. It is argued that if A_1 does in fact get 1 bushel of corn for 2 oz. of tobacco instead of the 5 which he would have been willing to pay, he obtains a surplus of 3 oz.

The same, in different degrees, is said to be true of the other buyers, except the marginal ones whose evaluations coincide with the ratio of exchange. In the same way, if there is only one buyer who buys several units, it will be the utility to him of the marginal unit which will determine one of the limits of the ratio. Each unit above the marginal unit possesses a higher utility, and being bought at the same ratio will produce for the buyer such a consumer's surplus.

It has been said that this argument overlooks an essential fact: the respective maximum and minimum offers of all these buyers and sellers simply meant that we assumed that they would have been willing to offer or demand these ratios had there been only one unit of the respective good available. It is an assumption which follows from the whole conception of relative marginal utility; and we have made use of it only as an aid in our analysis. It does not entitle us to conclude, so it is argued, that when the exchange is, in fact, made under conditions such as we have assumed there is any significant sense in which we can speak of the difference between what an individual would have paid under different conditions and what he has actually paid.

It is also argued that, since utility is not measurable, the concept of consumer's surplus is illegitimately derived from our concept of the scale of preference. It assumes that utility can be measured, which we have seen to be untrue. It regards the utility of a good as something absolute, whereas we have seen that the

term utility only has meaning as a concept expressing a relation in specific circumstances between the want for one good and that for others. It implies also that each unit of a stock of goods can have different utility, which we know is not the case.

While it is true that no quantitative precision attaches to the notion of consumer's surplus, it can be argued in its defence that it only implies that the existence of a good with a uniform ratio of exchange enables individuals to increase their aggregate satisfaction. The truth involved in the concept is that for any individual who buys more than one unit of it, the existence of a particular good at a certain ratio of exchange increases satisfaction. It is only on very special assumptions (which we cannot discuss here) that it is claimed that this increase in satisfaction can be measured.

4. MONOPOLY. PERFECT AND IMPERFECT COMPETITION

It is not only the number of buyers and sellers, but also the quantities for which they wish to contract which must be taken into account. What will happen if we assume the possibility of variations in numbers of buyers and sellers and quantities offered and demanded? If one seller appears with 7 bushels of corn, which he is willing to sell at $\frac{1}{4}$ oz. each, he will attract to himself all the buyers. Similarly if one of the buyers is prepared to buy 7 bushels at 5 oz. each, he will attract the whole supply and all the other buyers will have to remain unsatisfied. The state of affairs, then,

will be that of monopoly (either on the seller's or on the buyer's side) which was mentioned in the preceding section.

Monopoly. Thus, where there is a single person capable of controlling the whole amount to be exchanged, he will be able to determine the ratio; or, alternatively, if he fixed the ratio, the actions of the many people on the other side will determine the amount he can dispose of. The monopolist has, then, three possible policies. He can discriminate, that is, keep the people on the other side in separate compartments and strike separate bargains with each. This is the policy of discrimination which we have already met. He can, secondly, fix the total amount he is willing to sell and allow a uniform ratio to be determined by the scales of preferences of the buyers. And, thirdly, he can fix a ratio and allow the scales of preferences of the buyers to determine the amount which can be cleared at that ratio. In either of these two cases he would have control within the limits of the given scales of preferences of the buyers. The same is true, in a less degree, where one individual controls not the whole supply but the bulk of it, while the rest is in the hands of a large number of others. The policy which he will adopt will, of course, be that which will give him the best bargain. It is conceivable that it would pay him best to sell little at a high ratio or much at a low one, depending on different scales of preferences of buyers. We shall see later what are, broadly, the different possibilities.

Perfect competition. The opposite case is that where we have so large a number of participants that no single buyer controls an appreciable portion of the demand in relation to the total and no seller controls an appreciable proportion of the total supply. In this case we speak of *perfect competition*. We shall presently see this exchange relation in detail. But we can already say that, under perfect competition, the ratio will be independent of the demands and offers of each individual in the sense that each individual will regard the ratio as given and adjust his actions accordingly. This follows, notwithstanding the fact that the actions of all buyers and sellers bring about a certain ratio, from the assumptions that all the buyers and sellers are in communication with one another and that none of them controls a large proportion of the demand or supply. For example, in the case cited above, it is inconceivable that B_1 should be able to get from A_1 the maximum (5) which the latter would have been willing to pay if only one unit were available; for there are other sellers who would step in and offer their units at lower ratios. Nor, on the other hand, is it possible for A_1 to obtain his unit from B_1 for the minimum for which the latter might have been willing to part with it ($\frac{1}{4}$), for there are other buyers who would offer more. In this way, a ratio must emerge which is uniform for all the dealings that take place and at which the competition between buyers and sellers is in equilibrium.

The speed with which such a uniform ratio will be

established will vary with the concrete circumstances of the case. Sometimes lack of knowledge of the likely actions of others may lead a buyer or seller to postpone declaring his bid. Sometimes the bid may be erroneous when compared with the ultimate equilibrium ratio, and it may take a number of oscillations around the equilibrium, with consequent extra gains and losses, before the final ratio is established. In the case of some dealings, on the other hand, the intentions of the participants are so quickly revealed and made generally known that equilibrium is quickly reached. And once it is reached, it will form the basis of individual demand and supply.

Imperfect competition. Between the two extremes of perfect competition and monopoly there may be a number of variations. In these intermediate cases there is neither a single individual who controls the bulk of the amount demanded or supplied, nor are there so many that their individual shares are negligible in relation to the total. We find a number of people (buyers or sellers) who are each able to influence, in some measure, the terms of exchange.

These intermediate cases are, in many ways, the most important. First of all neither monopoly nor perfect competition are as typical of the conditions of exchange in reality as are these intermediate cases. (Although this is not to say that the analysis of these two extreme forms is not essential for an understanding of reality.) The intermediate cases present, moreover, particularly difficult theoretical problems which

have been the subject of much recent work. It is impossible, on this elementary level, to discuss them in detail, but some reference to them is necessary.

Duopoly. The situations intermediate between complete monopoly and perfect competition are generally described as *oligopoly*, and, in their simplest form as *duopoly*. The latter situation exists when there are two sellers (or buyers), each capable of influencing amount and ratio, and a sufficient number of competing buyers (or sellers) to make the laws of competition applicable to them. The determination of the terms of exchange under these conditions, i.e. the amount which each will be able to dispose of and the ratio which will rule, is one of the most controversial problems in economics, and there is, as yet, little agreement. There are, broadly, two schools of thought. One believes that duopoly can produce an equilibrium which will be stable and which will lie somewhere between the equilibrium of complete monopoly and that of perfect competition. The other school holds that, under duopoly, there is nothing in the conditions of exchange themselves which will produce equilibrium.

The first of these opinions goes back to the solution given last century by the French economist Cournot. According to him, any change in the offer of one duopolist will make the other adjust his offer so as to obtain the best bargain. In other words, he will regard the offer of his rival as an independent variable. The other duopolist will behave in the same way.

Readjustments will take place until each one reaches a point from which he has no incentive to depart, the offer of the other being unchanged. But, as Professor Pigou has recently shown, it is inconceivable that each rival can be aware of his ability to influence the ratio of exchange, which is inherent in the duopolistic situation, and yet believe that any change of his offer will not cause the offer of the other to vary, which is the condition for Cournot's equilibrium.

It has been suggested, therefore, that the situation is an essentially unstable one, for each rival will endeavour to attract a larger share to himself, if necessary, by an attempt to reduce the ratio. More recently, however, other weighty arguments have been adduced to show that a stable position can be reached. It has been suggested by Professor Chamberlin and others that both rivals will realize the futility of their original assumptions as to each other's reactions and will come to a 'tacit understanding'. In this way, or even by an open agreement, the situation will be transformed into one of complete monopoly with the two former rivals sharing in the monopolist's offer.

Against this it has been pointed out that even were this the result, it would only provide temporary stability; agreements may be renounced; and the 'tacit understanding' may, at any time, give way to a fresh attempt to drive the rival out. At any rate, an open agreement could not, strictly speaking, be considered as a solution of the duopoly situation, since it changes that position itself.

More important in practice than the case of two or a few sellers selling the same product is that of a number of sellers each selling a slightly differentiated product. We shall see later some of the implications of the attempts to differentiate similar products.

5. MONEY

Indirect exchange. From the results of our inquiry it should be possible to arrive at certain general rules which govern exchange. Before, however, we proceed to formulate these it is necessary to widen our assumptions in one important respect. We have hitherto been concerned with the case of two goods only. We must now consider the case of three or more commodities. As soon as three commodities are exchanged, the achievement of an equilibrium ratio becomes more difficult. Suppose, for example, that A is prepared to exchange 1 bushel of corn for 2 oz. of tobacco with B, who, in his turn, is prepared to exchange with C 1 oz. of tobacco for 3 pints of wine. A, who also wishes to obtain 1 pint of wine, finds that C is prepared to exchange it for $\frac{1}{2}$ bushel of corn. A's own evaluations, however, lead him to offer a maximum of $\frac{1}{3}$ bushel of corn for 1 pint of wine. So long as the three pairs remain isolated, A can exchange with B, B can exchange with C, but A and C cannot exchange.

At this point, however, A may realize that instead of offering corn to C in exchange for wine, he could offer him tobacco (which he can obtain from B). In this case the exchange relation will now appear as

follows: A gets 2 oz. of tobacco for 1 bushel of corn from B which he offers to C, from whom he can get as much as 6 pints of wine for it. In other words he can, in this way, obtain 1 pint of wine for $\frac{1}{6}$ bushel of corn instead of $\frac{1}{2}$, which was C's original demand and which A was not prepared to offer.

Exchange was made possible, in this example, by being made *indirect*. We can conceive of A not wanting any tobacco for its own sake at all; he may nevertheless endeavour to get some by exchanging corn for it, since he knows that tobacco will procure him the wine which he really wants. In this case, the marginal utility to him of tobacco will be affected through the indirect purpose which it can serve. While previously tobacco may not have appeared at all on his scale of preferences (because he is a non-smoker), yet, since it can serve to obtain wine, which is high on his scale of preferences, it will at once rise on the scale. The place it will occupy, in relation to that of the other goods, corn and wine, will depend on the ratios of exchange which can be obtained. It is conceivable, for example, that if B discovers that A might have been prepared to give $\frac{1}{3}$ bushel of corn for the pint of wine which now, with the aid of B's tobacco, he gets for $\frac{1}{6}$, B may put up his ratio of tobacco for corn. The possibility of indirect exchange will thus be one of the most important determinants of the scale of values; and we shall see, shortly, how the existence of habitually indirect exchange affects the scale of values of the participants.

The medium of exchange. In the example just

given tobacco played the part of a *medium of exchange*, that is, it facilitated the exchange of corn and wine between A and C. It was acquired by A not for its own sake, but for the sake of what could be bought with it. It requires little imagination to see that this process can be extended. C, for example, may not have wanted tobacco either, but knowing of another individual, D, who for tobacco was prepared to sell fish (which C wanted), he was, in his turn, ready to accept tobacco. Thus no difference in theory arises when we enlarge the circle of exchange by including more people and more goods. There are, however, important differences of a practical nature.

In our simplest case of three people and three goods, one good was chosen on the spot as a medium of exchange. The reason for the choice was the knowledge that it would be accepted in exchange. When the circle grows, this condition will be more difficult to fulfil. It is conceivable that, at first, any participant will accept a good in exchange if he has a reasonable hope of being able to pass it on and if, in addition, he has some use for it himself. He knows, then, that even if he does not succeed in finding the person to whom he can pass it on, he will yet have acquired something which possesses utility to himself. In the first place, therefore, the good chosen as a medium of exchange will have to be one the utility of which is derived not only from its expected ability to satisfy what we might call the want for exchangeability, but other wants as well.

Money. We give the name of *money* to a good which, in particular circumstances, is habitually chosen as the medium of exchange. In a comparatively restricted circle of exchange (such as we should find among the members of a primitive tribe), the conditions which the good will have to fulfil will not be very stringent. But as the circle grows, as more and more people with different scales of preferences and an ever greater variety of goods are drawn in, the conditions of acceptability to which money will have to answer will become more severe. Cattle, hides, minerals, many different goods have been used as money from time to time. They were chosen because at the time, and for the people concerned, they possessed a great value. Everybody who accepted them could feel reasonably sure that he could pass them on. In any case, they were goods which he would wish to have.

Gold. In the more complex exchange conditions of the modern world, money has almost universally been one of the precious metals, in particular, gold. Originally the qualities of lustre, durability, and rarity possessed by gold made it highly valued as an article of ornament and, no doubt, contributed to its adoption as a medium of exchange. But gradually these qualities fell into the background, and others which made it particularly efficient in its medium of exchange function became more important; and it acquired a conventional acceptability which was based partly on its physical qualities. These physical qualities are, strictly speaking, irrelevant; we only

need to know that gold does possess utility (for whatever reasons) in the same way as any other good; but they may be mentioned here.

Gold is homogeneous: one piece is the same as any other. It is very durable: the owner can feel that it will not deteriorate if he holds it in the expectation of some future act of exchange. It is easily transported, since it possesses great value in relation to bulk. It is easily divided into very small fractions which can be graded, weighed, and stamped. This, as we shall see again in a moment, is very important in enabling the bargaining partners to achieve a very fine adjustment of the ratio. There is one other characteristic of gold which is sometimes put forward in explanation of its adoption as money: the fact that, because of its durability and the comparative scarcity of its occurrence, additions to the already existing stock must be proportionately small. This means that the supply cannot vary greatly, and the value of gold will not be subject to great disturbances, at any rate in so far as it is determined by the quantity available. This is a very controversial point, and a full discussion of it goes beyond the scope of this book.

6. THE COMMUNAL SCALE

Our next question is, to what extent are the exchange relations, discussed in the preceding pages, affected by the existence of a generally accepted medium of exchange? To the remarks made in part I we must now add a further theoretical analysis.

The 'quantification' of subjective scales. We have met a certain difficulty earlier with the problem of the 'quantification' of subjective evaluations. We noted how, in spite of the immeasurability of utility, the scale of preferences obtained quantitative expression when the need arose to decide upon the worthwhileness of exchange. The difficulties of giving such quantitative expression to the scale of preferences are vastly increased when we consider the problems with which an individual is, in reality, faced. He has a complex structure of a large number of wants; he desires to achieve the maximum satisfaction; and to this end he will often have to rearrange his varied stocks of goods by means of exchange with a large number of different people. But if there is a generally accepted exchange medium, money, he will be able to express his willingness to exchange in terms of money. The same will be true of all other individuals. They, too, will refer their exchangeable goods to the common monetary denominator, and the monetary evaluations of each can be compared with those of the others. This, however, does not mean that a similar quantitative relation exists between the subjective evaluations themselves; we shall see in a moment what exactly the relation is.

The transformation of subjective scales into objective facts and their quantitative expression, which already takes place in any act of exchange, is made more easy when exchange is indirect, particularly when it is habitually monetary exchange. We have

seen that in exchange the tendency is to establish that ratio at which the relative marginal utilities are equal for the participants. In monetary exchange, this tendency persists and is generalized. The uniform exchange ratio, which we found will emerge in certain conditions, will, in indirect exchange, be expressed in money and it will express the above-mentioned identity of *relative* marginal utilities for all participants.

The objective scale. Wicksteed has expressed this in his concept of the *communal scale*. There will exist in the exchange system a communal scale which, in terms of money, expresses the relative marginal utilities of all exchangeable goods. This scale will be the same for all individuals with two qualifications only. In the first place, the communal scale will only relate to those goods which habitually enter into exchange. Those things which are impossible of monetary calculation, which men do not, in fact, freely exchange with one another, will not enter into the scale. Nevertheless, they too will be included in the subjective evaluations of each individual.

The second qualification is that the communal scale will be relevant to each individual only in certain parts, namely, those which contain the goods which he wishes to buy and sell. Alcohol and tobacco are freely bought and sold; they therefore occupy places on the communal scale. But to the abstainer and the non-smoker the portion of the scale which contains them will not be relevant.

We must examine a little more closely the way in

which that communal scale is arrived at, and the way in which the existence of such a communal scale affects the conduct of the individual.

Price. It is necessary, first, to re-define some of the terms used earlier and to generalize our exposition of the process. We call the *price* of a good the number of units of money which is given in exchange for one unit of it, i.e. its exchange value in terms of money; the *demand* for a good we now define as the quantity of that good which people are willing to buy at a certain price; and the *supply* is the quantity which people are willing to sell at a certain price.

The communal scale which is the result of the existence of exchange, the existence of money, and the interactions of many buyers and sellers will appear to each individual as an objective factor and, in accordance with it, he will organize his own scale of preferences. Monetary calculation leads, therefore, to rearrangement of subjective scales. A picture which I may possess may stand low in my scale in relation to some diamonds; but if I know that I can obtain for that picture a certain sum of money which will enable me to buy the books which stand higher in my scale even than the diamonds, the relation between picture and diamonds will be reversed. Or, as it is sometimes put, marginal utilities will be proportional to prices.

The more exchange is practised, the more will this rearrangement of scales proceed. The everyday conduct of individuals, in so far as it is concerned with

exchange, will relate to the objective scale. The relation between what is and what is not on the objective scale, and the extent to which that scale will be relevant, will naturally vary from time to time. In earlier times certain goods, such as land, were legally excluded from general exchange and could, therefore, not obtain a place on the objective scale. Goods that are rare and are only infrequently the object of exchange (certain precious stones, or old masters, for example) do not appear on the objective scale at all; though for many even of these, exchanges are now sufficiently frequent to result in a price.

In modern communities the number of goods on the objective scale has been much increased, since most of the legal obstacles to their exchange have been removed. There are still, however, many goods that one does not, normally, regard as exchangeable, and to which one does not, therefore, attach a price, such as love, friendship, and honour. It is conceivable that these, too, could form the object of exchange: even to-day, in many countries, the favour of officials and ministers has its price.

Nor will the same individual always have the same idea of the relevance to him of the objective scale. I may possess furniture, books, pictures, to which in the ordinary course of events I should never attach a price (once I had acquired them), although, being frequently dealt in, they appear on the objective scale. But if I fall on hard times and have to consider rearranging my affairs, I may once again regard them

as exchangeable objects, and the communal scale relating to second-hand furniture, books, and pictures will then have relevance to me.

The value of money. It may be useful, at this point, to mention the existence of a special problem of the value of money, even though a detailed discussion of it is beyond our purpose.

This problem is of importance, since any change in the value of money will mean a change in the general relation in which money and other goods will exchange for one another. Any alteration in the willingness of people to hold money as against other goods must affect the supplies of and the demands for goods and their prices. In a sense, of course, the problem of changes in the value of money is not a special one. It must be governed by the laws to which relations of exchange in general are subject. In particular, we could apply to it the analysis which will be used later in the case of goods with correlated utilities (complements and substitutes) and the concept of demand elasticity.

The treatment of the problem has, however, constituted a special department of economic theory, because the relation between monetary changes and price changes is, of necessity, a particularly striking one. Certainly, during the last 150 years there have been many price changes of so violent a character that an explanation which considers only changes in the supply and demand conditions of different goods becomes impossible. When changes occur which seem to affect the prices of all goods in a similar manner, the

presence of a monetary influence can rightly be suspected.

The quantity theory. One of the earliest statements of the relation between money and prices is embodied in the *Quantity Theory of Money*. Briefly, this theory states that since the value of money, like the value of any other good, is the amount of other goods for which a unit of it will exchange, its value, other things being equal, will depend upon its quantity in relation to the quantity of other goods. Its value will tend to move inversely to any movement in its quantity. Thus, an increase in the amount of money which is not accompanied by an increase in the volume of goods to be exchanged will tend to raise prices in general. This is sometimes expressed in mathematical formulae as, for example,

$$\frac{M}{T} = P,$$

where M is the effective amount of the circulating medium (including all kinds of money and money substitutes multiplied by the number of times each unit of money performs its function in the period under consideration—what has been called the velocity of circulation), T is the volume of trade carried on, and P is the general level of prices.

There can be no doubt that the Quantity Theory expresses an important relation. But this equation of exchange, in its usual form, is of so general a character as to be nothing more than a truism. It is only a statement to the effect that the quantities of goods

exchanged, multiplied by their average price, equal the volume of money multiplied by its velocity of circulation. In other words, it is only a crude statement of the interrelation of supply, demand, and price.

To make it yield more, a much more detailed inquiry is necessary. The aggregates (volume of money and volume of trade) and the average (the general level of prices) of which it is composed must be broken up into more concrete concepts. An examination is required both of the conditions which determine changes in the demand for and the supply of money and of the effects of such changes on different kinds of prices, since not all prices are uniformly or simultaneously affected by monetary changes.

It is impossible to study recent attempts in these directions at any length here: the theory of money still remains, for didactic purposes, at any rate, a special branch of economics. But some indication may be given of at least one recent trend which shows that the substitutional concepts used in the treatment of the general theory of exchange are being applied to the theory of money.

The demand for money. In the sense in which we have described it earlier, money has no value of its own beyond that which it derives from its non-monetary uses. It is simply the unit of account with the aid of which comparisons are possible between the value of one good and that of all other goods, whenever an act of exchange is contemplated. On these occasions it can be said to represent value. The

position is, however, changed whenever the question arises whether to hold resources in the form of money or not. It becomes pertinent to inquire into the value of money. For then the individual will be balancing the advantage of holding goods against those of holding money, and will thus assign to money a place on his scale of preferences. But the individual will only be in a position to have to decide whether to hold money or not if the acts of exchange are spread over time. If we could conceive of money being employed as a medium of exchange in the sense used earlier, it would never be in anybody's possession for any appreciable time; it would be constantly changing hands and only the relative values of the goods themselves would be of importance.

We know, as an elementary fact of experience, that the future is, in some degree, taken into account by the individual in his scale of preferences. But even when an individual considers future exchanges it can be shown that he will not necessarily hold money. Since money possesses no direct utility of its own, it will always be more advantageous to hold goods. If these future exchanges are certain as to time and terms, he can so arrange his dealings in general that he will be prepared for them, and we shall see later that the possibility exists of exchanging present and future satisfactions. It has, therefore, been suggested (most recently by Dr. Hicks) that it is only due to the *uncertainty* of the future that a demand for money, as such, exists. It is uncertainty of the future which

creates a problem of the value of money, as distinct from the values which money represents in the acts of exchange.

In the further examination of this problem it becomes necessary to inquire into the different factors which, by affecting individuals' anticipations of the future, will affect their demand for money. Similarly, it is essential to examine carefully the 'path' which any increase in the supply of money takes, i.e. the sequence in which it affects different groups of individuals and thus alters the terms on which they are willing to hold money or goods.

These questions cannot, however, be pursued here; they are mentioned in order to direct the attention of the reader to the more advanced problems of our subject.

7. THE LAWS OF EXCHANGE

For purposes of convenience we must now make certain assumptions. We will consider the individuals who are effecting exchange to have given stocks of goods which they exchange with the aid of a medium. To make the exposition simpler, we contemplate each individual as exchanging one good only for money. We have already seen how, with a multiplicity of buyers and sellers who are in communication with one another, a uniform price will be established at which all exchanges will take place, and how, with the aid of money, a communal scale will result. We can now state it in this way. If, at any given price, a certain quantity of a good is bought, the marginal

utility of that good is diminished in relation to the marginal utilities of all other goods (represented, in this case, by money). The more of the good is bought, the more, relatively, will its marginal utility decline.

The sellers we can now see to be in exactly the same position as the buyers, i.e. people who exercise a certain demand for the good. It is to be noticed, therefore, that the distinction, at this stage of our theory, between demand and supply, and between buyers and sellers, is not really necessary. The sellers, too, demand the good which is exchanged; but their demand for it bears, so long as exchange continues, the opposite relation to their demand for other goods (money) as does that of the buyers. When exchange is completed the ratios are the same. As it is, however, customary to speak of supply and demand, we will continue to use both terms, warning the reader, once again, against treating them as different in kind.

The demand schedule. It follows from the above that the higher the price of a good the less will be demanded. Suppose an individual demands a certain quantity at one price, that means that at that price he is prepared to part with a certain amount of money (other goods). If the price rises he may be prepared to part with as much money (other goods) as before, or even with more, but he will not, normally, spend so much more money as to buy a larger quantity of the good. If he did that, he would be buying units of the goods possessing less marginal utility than other goods which he could buy with that additional expenditure.

It follows, therefore, that demand for a good will be inversely related to price. While this holds good for buyers and for sellers, it is more customary to say that supply is directly related to price.

We can draw up a schedule showing this relation between demand and price, and if we add the schedules of all demanders we obtain a schedule for the community as a whole. The price will then be uniform for all and, at that price, the whole stock will be cleared, provided that we include the demand of the sellers, at that price, as well. Or, as it is more commonly expressed, and as was shown in preceding sections, at that price supply and demand will be equated.

The above analysis has shown us, on the one hand, how the communal scale is fixed by the actions of all participating in exchange. On the other hand, it has demonstrated that each individual will arrange the disposition of his own stock of goods, in so far as it is available for exchange purposes, in accordance with that communal scale. Individual demands (and offers) and prices are thus mutually dependent. It is easy to see that this interdependence extends to all goods which an individual demands, since the price he is willing to pay for a good is only the expression of the relation between his demand for that good and his demand for all other goods.

The elasticity of demand. There are only a few points to be added before this analysis is completed. The conclusion we came to, that demand is inversely related to price, enables us to adopt the concept of

elasticity of demand. This term expresses the degree of correlation between demand and price. If a rise in price produces a marked fall in demand (and a fall in price a marked rise in demand), we regard the demand as elastic. If a marked rise in price produces only a small decline in demand (and a marked fall in price only a small rise in demand), we speak of an inelastic demand. Or, to be more precise, we can distinguish between demand elasticities which are unity, less than unity, and more than unity. Demand elasticity is unity when a change in price produces a proportionate change in demand; it is less than unity when the change in demand is less than in proportion to a change in price; and it is more than unity when a change in price calls forth a more than proportionate change in demand.

With a demand elasticity of unity, the total outlay on a good in relation to others will remain the same at all prices. With an elastic demand a fall in price will produce a greater outlay, and a rise in price a smaller outlay. Finally, in the case of an inelastic demand, a fall in price will result in a smaller, and a rise in price in a greater, outlay. It is to be understood, however, that our general rule must hold; an individual will not, normally, buy at a higher price a greater quantity than he was prepared to buy at a lower, or a smaller quantity at a lower than he was prepared to buy at a higher price.

It is not possible to lay down a general rule as to the demand elasticities of particular goods. Sometimes

it is pointed out that the demand for the necessities of life is inelastic, while the demand for luxuries is elastic. This is only true as a generalization for a particular time and place. It is true that there are always some goods for which some individuals will pay a high price if they cannot obtain them at a lower price. There is a minimum of food, shelter, clothing which we must have; and we would, if necessary, part with the whole of our limited resources in order to obtain them. The comforts of life, on the other hand, we should more willingly be prepared to give up if their price rose and if to acquire them placed too heavy a burden on our limited resources.

But it is never easy to say what particular goods are included in the two categories. The idea of what is and what is not a necessity varies notoriously for different individuals according to their tastes and according to the total resources which they possess. And even the same individual will change his demand elasticities, when, from time to time, his tastes and his resources alter.

Monopoly and competition re-examined. The concept of demand elasticity is an extremely valuable one, for it enables us to express more precisely some of the exchange relations previously studied. In particular, we can now re-formulate the different ways in which price will be determined under monopoly and competition.

In the case of monopoly we have seen that if the supply is fixed the price will be fixed also, and vice

circumstance tending in that direction would be anything which produces the belief in the mind of the buyer that the goods of one seller are not exactly the same as those of any other seller, provided that he acts on that belief. In that case, that particular seller would be able to depart, after careful calculation, to a certain extent from the uniform competitive price. This differentiation of goods may also provide the most realistic solution of the duopoly problem previously touched on, for it would help to create, as it were, a non-competitive area for each rival.

8. COMPLEMENTS AND SUBSTITUTES

There is a further special point which it is important to consider. At an earlier stage we noticed, among the characteristics of wants, their competitive or alternative nature and their complementary character. To these correspond certain characteristics of goods. It is customary to distinguish between *substitutes* and *complementary goods* (or, as Professor Irving Fisher has called them, competing and completing goods). We speak of substitutes when two goods are so related that either one or the other may satisfy a want. Two goods are complementary, on the other hand, when the want for one good involves also the want for the other. One speaks generally of *independent* goods when neither of these two relations exists.

It is clear that when relations of the above nature are present, the marginal utilities of the two goods will be peculiarly related. Their demands and prices will,

therefore, also show a special correlation which requires consideration. Recently, however, it has been shown, by Mr. R. G. D. Allen and others, that the above definitions are not quite precise. In the first place, in a general way, all goods may, from time to time, be either substitutes or complements depending on the tastes and resources of the individual at the particular time. Moreover, each good when considered as part of a structure of goods (which forms a certain standard of living) is always a complement; while when we look upon goods as competing for a limited amount of resources (money), all those within a given group, e.g. food, are substitutes.

The relation of marginal utilities. We can, however, establish the following general rules. The *normal* relationship between any two goods is such that an increase in the quantity of that good reduces its marginal utility in relation to that of the other good. A *non-normal* position exists when one good is superior and the other inferior; in this case an increase in the quantity of one good always reduces the marginal utility of the inferior good in relation to that of the superior good.

Goods in the normal position are all goods generally regarded as independent as well as all complements. An increase in the amount of X will diminish its marginal utility in relation to that of Y (or increase the marginal utility of Y in relation to that of X); and the converse is true of any diminution in the quantity of X. This applies equally to motor-cars and wireless

sets, and to motor-cars and petrol. The only difference is that in the case of the latter, which is more generally regarded as a complementary relationship, the change in the marginal utility relation will be very rapid, while in the former case it will be slower. Thus an increase in the quantity of motor-cars will raise the marginal utility of petrol in relation to that of motor-cars quickly and markedly, but that of wireless sets more slowly and less markedly. Clearly, many different grades of correlation are here possible according to the closeness with which the two goods are bound together.

Goods will be in the non-normal position particularly often when the relative quantities of the two goods are very uneven. Suppose, for example, that an individual possesses a large quantity of bread and a small quantity of meat. Meat is the superior, bread the inferior good, because, when possible, the individual will substitute meat for bread. An increase in the amount of bread will, as in the previous case of complements, diminish the marginal utility of bread in relation to that of meat. But in this case an increase in the amount of meat, the superior good, will diminish the marginal utility of bread even more quickly than that of meat. In the case of an increase of either good, therefore, there is a relative diminution in the marginal utility of the inferior good (the substitute, more commonly so called).

The effect on the price-demand relationship is obvious. In the case of complements, a fall in the

price of one good will increase the demand for the complement in relation to the demand for it, and vice versa. In the case of substitutes a fall in the price of the superior good will increase the demand for it in relation to the demand for the inferior good. A fall in the price of the inferior good will diminish the demand for it in relation to the demand for the superior good. Finally, where a third, non-related, good falls in price, the demand for it will rise, but if the elasticity of the demand for it is less than unity, the demand for the inferior of a pair of substitutes will fall in relation to that of the superior good.

All these statements are true only with the proviso—which we have tacitly made all through the present analysis—that the total outlay for all goods remains the same. Should the total outlay rise, then the demand will rise, but not uniformly for all goods. The demand for the superior goods among substitutes will rise, as will the demand for all other goods. But the demand for the inferior goods will fall.

All these cases are of very great importance in practice. Not so long ago, for example, one could study the relation between butter and margarine, a superior and an inferior good. The low price of butter led to an increase in the demand for it and to a substitution of it for margarine. The consequent relative fall in the demand for margarine led to a fall in its price also.

When the total outlay of a group of people declines, as, for example, in a depression, the general tendency

is for the demand for the inferior good to increase in relation to the demand for the superior good, and to all other goods. The opposite is true of a time of prosperity, when total outlays increase.

Giffen's anomaly. One abnormal case, first described by Sir Robert Giffen, might also be mentioned. A rise in the price of the inferior good will, as we know, diminish relatively the demand for the superior good. Sometimes this may be brought about by an absolute rise in the demand for the inferior good as well as by a fall in the demand for the superior good. In this case, we would have an exception to the normal rule that an individual will never increase his demand for a good when its price has gone up, even though he may increase his total outlay. This exceptional state of affairs arises, according to Giffen, in the demands of the poorer classes for meat and bread. A rise in the price of bread will, on account of the small total outlay, the urgency of the need for the good, and the marked superiority-inferiority relation between the two, lead to a rise in the demand for bread, and a greater substitution of it for meat. From the above analysis it will be clear that this exception from the norm can be easily explained.

9. EQUILIBRIUM

We have just noted one assumption which was made throughout the preceding analysis of exchange: that of a given total outlay for each participant. This, however, is not the only assumption which conditions

the validity of the laws of exchange at which we have arrived. It is necessary here to recapitulate the assumptions made explicitly, and to state those which were implied in the argument.

In all the cases treated above we have assumed a number (at least 2) of people with given amounts of goods, actuated by the desire to achieve as nearly as possible a uniform level of marginal utilities in accordance with the second law of Gossen, led, thereby, to exchange, and using money to facilitate exchange. In postulating this we have made certain assumptions as to the behaviour of the individuals with whom we have been concerned and as to the quality of their interrelationship.

The market. We have assumed that the conduct of our individuals is in accordance with the laws of choice, that they are free to act as they choose, and to contract with one another freely. Some of the implications of these postulates we shall discuss; for the moment it is clear that our individuals must be in communication with one another; i.e. there must be the physical possibility of exchange and there must be knowledge on the part of each participant of the actions of the others. In such a case we say that the buyers and sellers meet in a *market*.

In the ordinary sense of the word, a market implies a physical meeting-place of buyers and sellers, such as is found in the small local market-place to which potential buyers bring their demands and in which they discover the offers of sellers. But in our usage the

term market denotes something much wider than this. It describes the machinery by which individual subjective evaluations are transformed into the objective scale of prices; and the market for any individual good performs the function of establishing equilibrium between all the subjective evaluations of that good. That equilibrium will, as we know, be reached when the marginal utility of the good in relation to that of other goods is the same for all who possess it and when it is higher for those who have it than for those who have not.

When we speak of the world wheat market, we do not have in mind any physical place (although there are important exchanges in which buyers and sellers of wheat meet); we mean the fact that buyers and sellers of wheat all over the world are in constant communication with one another, and that beyond the physical market-places there is a mechanism of communication which unites them all. Similarly, when we say that there is a market for X, but that there is no market for Y, we mean that there are frequent dealings in X which have put it on the communal scale, and have given it a price; and although there may occasionally take place exchanges of Y, these are so infrequent that there is no price for Y.

Market equilibrium. In a market in which buyers and sellers are in perfect communication, there exist certain relationships between demand, supply, and price which we have studied. We can summarize our results by saying that demand, supply, and price

are mutually dependent. Demand depends on price; supply depends on price; and price, in its turn, is determined by the relation between supply and demand; or, to put it somewhat differently in accordance with the statement that it is not necessary to speak of supply at all, the demand of each individual will depend on price, and price is determined by the interrelations of all demands. We have, moreover, seen that the self-regulating action of this system of interdependence is such that any movement of any one of the factors tends to produce an opposite movement. We can, therefore, conceive (in theory) of one relationship between supply and demand at which no change in price will occur. Whenever in the preceding pages we have studied the determination of prices, the demand and supply conditions assumed have represented such a relationship.

This is true of each market (i.e. the mechanism of exchange of each individual good). But there is also a relation between all markets: owing to the general limitation of all resources, the demands for and the supplies of all goods are interrelated. We can also conceive of such a relation, or system, of all prices that there will be no tendency for any demand and supply relations to change. To both these conditions we may give the name of *equilibrium*, the former applying to the exchange of a single good, the latter to exchange in general. At the conclusion of the theory of production we shall return to the concept of equilibrium. We shall have to discuss again, with wider terms of

reference, the conditions of equilibrium in general; the different kinds of equilibrium that are possible (i.e. under competition and under monopoly); and, in particular, the relation between this analysis and reality, i.e. whether, and, if so, under what conditions, the position described as one of equilibrium can be realized in practice.

III. THE THEORY OF PRODUCTION

I. THE SCOPE OF THE THEORY

IN the theory of consumption we examined the individual's disposal of a stock of goods. We then proceeded to assume that his acts of choice would extend to exchanging his goods with those of others. Here again, we advanced from the case of two individuals to the most general case of the competitive market and system of markets. All the time, however, we have had one qualification in mind; we did not inquire into the origin of the goods which were being exchanged, and assumed that the stocks in the possession of different individuals were fixed (notwithstanding that, from time to time, the quantities offered and demanded by them might vary).

In the theory of production, we assume that the physical stocks of goods can be varied and we study the implications of choice under this assumption. Productive activity includes everything directed towards making more means of consumption (goods of the first order) available to the consumer. The simpler of these activities (those which change only the legal availability of goods) we have already treated in the theory of exchange. It now remains to study activities which alter the physical characteristics of goods and what Professor Robbins has called their time and space indices. One must avoid any misunderstanding as to the relation between economics and technology.

Production, like all economic activity, involves certain technical problems. Technology shows how different means may be used for the satisfaction of different wants; in the theory of production it shows, in particular, how the physical and other characteristics of goods may be altered. But when the means for achieving many of these ends are limited and capable of alternative uses, an economic problem arises. It then becomes necessary to choose between the different uses to which these means may be put.

The original factors. The wants which human beings experience are directed to goods of the first order, i.e. those which are adapted for consumption. The extent to which such goods are available is considerably limited. But nature does provide certain goods which, when adapted, will become suitable for consumption. These goods we call the *original factors of production*. They include all the resources of nature to which we give the general name of *land*, and human energy to which we give the general name of *labour*. Production consists of the combination of these personal and impersonal factors for the purpose of providing want-satisfying goods.

The first thing to note about the factors of production is that they, too, are economic goods. They are scarce relative to the want for them (i.e. for the purposes of production as defined above). The second noteworthy characteristic is that the possibilities of productive combination of the factors are many, in other words they can be used in alternative ways. The

general laws of choice will therefore apply to them. They will possess utility and value; and, given the institutional conditions described in Part I, we can expect them to become objects of exchange. In that case, each will be dealt in in its own special market and will command a price.

We conclude, then, that the problem of the disposition of scarce factors by an isolated economy must be subject to the same rules as those found in the theory of consumption, while the exchange of scarce factors must be subject to the general laws of exchange. In a sense, therefore, the propositions already developed must contain generalizations capable of application to the factors of production.

The interdependence of prices. The reason why they are not sufficient as they stand but require further elaboration is that the demand for factors is not direct but is a *derived demand*. True, all demand is derived in the sense that we want means for certain ends, whatever these ends may be; but the indirect nature of the demand for factors is particularly obvious and itself requires explanation.

We can conceive of our economic man to be confronted, on the one hand, by a demand for goods of the first order, on the other, by a supply of scarce factors. The process of production has, somehow, to bring these two together. The theory of production has to show how these two are, in fact, brought together; how the demand for consumption goods is translated into a demand for factors, and how the

supply of factors is transformed into a supply of consumable goods.

We have already studied the implications of the demand for goods of the first order, in the pricing process. We have seen that, normally, demand and supply will, respectively, be inversely and directly correlated to price; and in the market there will, normally, be only one price at which demand and supply are equal. In addition, there is an interrelation of all prices resulting in a structure of mutually dependent prices. This interdependence of the prices of goods of the first order implies that the price of a good will be determined not only by the supply and demand conditions of its own market, but also by those of others. Even where, in a market considered in isolation, conditions may exist for the formation of a price at which supply and demand are equated, changes of prices in other markets will influence the price in our first market and thereby cause changes in supply and demand conditions.

Such an interrelation of prices is present not only between the consumption goods themselves, but also between consumption goods and factors of production. In studying this interdependence of prices, the theory of production analyses simultaneously the pricing of the factors of production and the quantities of factors used.

A division of the theory of production. There are two broad divisions in the theory of production. In the first, we ignore the possibility of indirect production

and study only production which is assumed not to take time but to result in the direct transformation of factors into consumption goods. This the Germans call *Momentanproduktion* (instantaneous production) and in English this branch of the theory is often referred to as non-capitalistic production. In the second part of the theory, we study the implications of indirect production which involves a lapse of time between the first use of productive factors and the appearance of consumption goods. This is capitalistic production. The reader's attention is, however, again directed to the proviso about the use of these terms given in Part I.

The first branch of the theory is again divided into two broad sections. The first studies production with the aid of a *single factor*; the second production by means of a *combination of factors*. Both these can, then, be treated for both the isolated and the exchange economy.

2. PRODUCTION IN THE ISOLATED ECONOMY

The case of single-factor production is the simplest, and we will therefore take it first.

The technical coefficient. We will assume that the individual possesses a given stock of a single factor, and that there are two possible uses to which that factor can be put. The individual will then have to decide between these alternatives. A scale of preferences will thus be established as an additional datum. There is yet a third datum which is necessary before

the uses of the stock of factors can be considered determinate, namely, the relation between quantity of factors and quantity of product. This is known as the technical coefficient. It expresses for each good the quantity of factors necessary to produce a unit of that good.

Suppose, for example, that the individual had 5 hours of labour (stock of factors), that it took 1 hour to produce 1 basketful of strawberries and 1 hour to produce 1 bagful of firewood (the technical coefficients), and that the scale of preferences as between strawberries and firewood was $1 \text{ basket} > 1 \text{ bag} > 1 \text{ basket} > 1 \text{ bag} > 1 \text{ basket}$. (In other words, the first basket of strawberries stands higher, but the second basket stands lower on the scale than the first bag, and so on.) If these are the data, it is clear that if a maximum of satisfaction is to be attained, then, according to the second law of Gossen, the 5 hours will be used to produce 3 baskets of strawberries and 2 bags of firewood. With this result, 1 hour of labour (one unit of the factor) will be producing the same utility, at the margin, in both uses.

The expenditure of effort. If the supply of labour is variable, the choice will be not only as to the uses to which it is to be put, but also as to the amount of labour to be expended altogether. It used to be argued that it was a special characteristic of labour that its use was accompanied by a subjective sensation which was the opposite of satisfaction. The expenditure of effort, it was said, produced pleasure, but it

also caused pain, and the two had to be balanced against one another. In more recent theory the formulation was to the effect that while the result of effort was utility, effort also gave rise to disutility. With the progress of formalism in economics it is difficult to see how such a distinction can be preserved. The 'pain' of labour can be regarded simply in the light of an increasing marginal utility of leisure as the amount of it declines in relation to the amount of the goods produced by labour. Leisure would thus occupy a place on the scale of preferences equally with other goods.

The scale of preferences, including leisure, and the technical coefficients will between them determine both the total amount of labour expended and the amounts of each good to be produced. If, for example, our scale of preferences as regards strawberries and firewood, and the technical coefficients were the same as before, and if leisure after 5 hours stood higher than the addition of a third bagful of wood, the individual would work 5 hours altogether and produce, in that time, 3 baskets of strawberries and 2 bags of firewood. In the older terminology it would have been said that the marginal disutility of 6 hours was greater than the marginal utility of 3 baskets plus 3 bags, but that the marginal utility of 3 baskets plus 2 bags was greater than the marginal disutility of 5 hours' labour.

Two factors and fixed technical coefficients. Little is altered if we assume the individual to possess two factors. Here, again, he will choose those uses which will give maximum satisfaction; and the scale of

preferences, the stocks of factors (where one factor is labour, the value of leisure will also appear on the scale), and the technical coefficients will, between them, determine the proportions in which the factors are used and the total quantity of goods produced.

Suppose, for example, that the individual possesses 4 units of labour and 4 units of a factor which we shall call land. He can use these to produce either good A or good B. The technical coefficients are 1 unit of labour plus 1 unit of land for each unit of either good. In this case, the total stock could be used to turn out a variety of combinations of these goods, i.e. 4 units of either A or B; 3 of A and 1 of B and vice versa; or 2 of each. In all these cases the total stock of 4 labour and 4 land would be used; and, in addition, there are all the productive combinations in which something less than the total quantity of factors is used. Now if the scale of preferences is, as before, $A_1 > B_1 > A_2 > B_2 > A_3 > \dots$ it is easy to see that maximum satisfaction will be obtained by using all the factors to produce 2 units of A and 2 units of B. An addition of a third unit of A could only be made by sacrificing the second B; and since the second unit of B stands higher on the scale than the third unit of A, there would not be equilibrium in accordance with the second law of Gossen.

3. PRODUCTION IN THE EXCHANGE ECONOMY

The institutional data. We now pass on to consider production in the case of an exchange economy. It is pertinent, at first, to consider briefly why exchange of

factors of production should arise. To ask that is to ask the reason for the existence of production as a social process. The reason for the exchange of goods of the first order was to be found in the fact that there existed differences in the ratios of marginal utilities between different people. It must, therefore, also be true that the individuals who own factors of production will value them differently and will wish to achieve equilibrium by exchanging them. Thus, a different subjective evaluation of factors of production is in theory the necessary and sufficient reason for exchange to arise. But differences in these subjective evaluations will, in fact, only arise because of certain specific objective conditions. These conditions have already been described in the first part of this book; they are *division of labour* and *co-operative production*.

The reason for division of labour is its greater productivity; it reduces the technical coefficients. This is a technical fact and is to be accepted by us as a datum. In its early stages, division of labour involved exchange of products. Each individual specialized in the production of one good, and while he ceased to be self-supporting, he obtained at the same time the possibility of procuring for himself the things he needed by the exchange of his own surplus products for those of others. We observe already, at this stage, that the individual owner of the factors of production does not use them for himself in the narrow sense, but only in so far as he can administer them in accordance with the communal scale.

When division of labour has developed to the point where each individual contributes only a small part to the production even of one single good, the problem of combining all individuals in co-operative production arises. A study of the way in which this problem is solved belongs to economic history. For our purpose it becomes a social datum. Co-operation in production is brought about to-day by the activity of private enterprise. It consists, briefly, of this. The owners of the factors of production offer them in a market in which the entrepreneur exercises a demand. The latter buys them, combines them in production, and sells their products in the market of finished goods. The owners of the factors appear again, as a body, in this market, but this time as buyers.

Co-operative production demands, therefore, a mechanism by which the two sets of markets may be connected. The entrepreneur acts as the connecting link; and while other such links might conceivably exist, it is through a study of the actions of the entrepreneur that we shall be able to answer the questions we set ourselves at the beginning of this part as to the relation between the market for goods of the first order and the market for factors in our social system.

The supply of labour. We have already discovered the general laws of the market for goods of the first order. Let us see how these laws apply to the market for factors now that we know that such a market exists: the demand comes from the entrepreneur, the supply from the owners of the factors. All we have to assume

at the moment is that the supply of factors will be related to the price. Sometimes the supply will rise (or fall) quickly in response to a rise (or fall) in price; sometimes it will respond slowly. We shall see later what special problems arise in connexion with the variability of the supply of factors, but something at least must be said here in relation to one factor, labour.

In general, there is no reason to suppose that the laws of price do not apply to labour. The labourer sells his services; he obtains a price which we call his *wages*; these, in a complex economy, are generally paid in money; he uses this money in order to buy consumption goods. In determining the terms of his offer he will have to balance the advantages he obtains against the disadvantages of the hiring out of his labour. Such balance must follow the general rule which relates demand, supply, and price. At a higher price, other things being equal, he will supply more, at a lower price, less.

It is, however, necessary to make a few qualifying remarks here; and more will follow later. In the first place, the seller of labour is in a different position from that of the seller in general. In a modern community the labourer is, as a rule, unable to obtain the goods which he wants by any other means than by hiring himself out. Even if he were prepared to become self-supporting—little though he could produce for himself—it is hardly possible for him to obtain the original impersonal factors which would enable him to do so. It is no exaggeration to say that unless he can find a

buyer for his services, he would be unable to maintain himself.

Now this is a social fact. But it has an important bearing on the relation between wages and the labour supply. The labourer will seldom be in a position to withhold his supply when the price falls; he will have no reserve price.

There is a further point: in certain circumstances, the supply of labour may react to changes in price in a directly contrary way to that postulated as the general rule. In other words, it may be that, below a certain point, a fall in wages will produce an increased supply; and above a certain point a rise in wages will produce a fall in the supply of labour.

The former of these exceptions follows from the first qualification made above. Where the labourer is anxious to maintain a certain minimum standard of subsistence, a fall in wages which threatens that standard will cause him to increase the supply of labour. He may work longer, or members of his family, who previously were not in work, may offer their services in the market. On the other hand, where wages are high enough to make the maintenance of a certain standard secure, a rise in wages may lead the labourer to curtail his supply and so increase his leisure.

4. SINGLE-FACTOR PRODUCTION

After these preliminary remarks, we may now proceed to consider the simplest case: that where the product involves the use of a single factor. We have

three sets of conditions to consider: those in the market for the product, those in that for the factor, and their connexion through the entrepreneur, who supplies in one market and demands in the other.

An example. The state of affairs is best illustrated by an example. We will use, with adaptations, one recently given by Professor Strigl. It must be made clear, however, that no example of the simplest case can be realistic, since we have to assume (*a*) the existence of two markets with entrepreneurs connecting them, (*b*) the use of only a single factor, (*c*) production to be instantaneous, i.e. prices are received by the entrepreneurs and are simultaneously paid out by them. These conditions are hardly likely to be found in practice. In particular, the use of only a single factor is seldom likely to go hand in hand with the existence of entrepreneurs in this line of production: it is more likely that the owners of the factors of production will themselves sell their products. The following example should be regarded as only a preliminary introduction to the problem of price and cost. More will follow later.

We assume, then, that there are entrepreneurs who deal in the service of window-cleaning. On the one hand, they have to meet a demand from householders for these services, on the other, they are faced with a supply of the services of window-cleaners. The demand of the householders will be inversely related to price. The higher the price, the less will be demanded: all householders will have their windows cleaned less

frequently; or some householders will clean their own windows; or a mixture of both these tendencies will bring about a diminution in the demand. Disregarding the exceptions mentioned in the previous section, we can assume the supply of the labourers to be directly related to the price. At a lower price, some will withdraw from the market and some will work less; while in order to obtain a greater supply, the entrepreneurs will have to pay a higher price.

It is clear that the entrepreneurs will arrange their demand in the labour market in accordance with the prices ruling, or expected to be ruling, in the market for the product—in this case the market for window-cleaning. They will not offer wages which they do not expect to be covered by the price which they can obtain for the service. Alternatively, they will endeavour not to sell any services at a price which is below the *cost* incurred for it, i.e. the wages of the labourers. We have here ignored the remuneration of the entrepreneur himself, since this makes little difference to the argument. We can simply reformulate the above statement and say that the entrepreneurs will be anxious not to sell at a price which is below cost including their own remuneration.

The law of cost. The mechanism which achieves these objects is known as the *law of cost*. It can be expressed in the following way. If the price is above cost, the entrepreneurs will obtain a large remuneration. Competition among them will, then, produce an increase in the supply, through each one

offering more of what is to him a very remunerative service and through a tendency for other entrepreneurs to enter this line of production. This increase in supply will lower the price. In the second place, the increase in the supply can only be achieved by drawing more labour into production, and this implies the offer of higher prices: wages will rise. Thus we have two tendencies, a fall in price and a rise in cost which cut into the remuneration of the entrepreneurs and bring about equality between cost and price.

Conversely, if price is below cost, the remuneration of the entrepreneurs becomes a negative one. The losses incurred in the window-cleaning business result in a diminution of the supply: some restrict their offers, others go out of business altogether; and this will tend to raise the price. In reducing their supply, the entrepreneurs also reduce their demand for labour and wages will fall. These two movements will go on until price and cost are equated again.

At what point exactly the two tendencies will meet cannot be determined. It will depend upon the scales of preferences in the two connected markets, i.e. on the elasticity of the demand for window-cleaning and on the elasticity of the supply of labour (on the reserve prices of the workers). It is possible, for example, that if the demand for window-cleaning is highly elastic, and if the reserve prices of the workers are very low, the meeting-point of the two forces will be at a low price-cost level. On the other hand, if the demand for window-cleaning is very inelastic, and if the

reserve prices of the workers are high, a high price-cost level will result.

The level will also be affected by the degree of competition among entrepreneurs. Indeed, without a consideration of the relation between individual entrepreneurs in the same line of production, little of real importance can be said about costs and prices. This will become clear when we examine the more realistic conditions of complex production. We shall see shortly that a more rigorous distinction between different kinds of cost is required for any exhaustive analysis. For the moment, the law of cost, in its simplest form, shows that the connexion between the product and factor markets in certain conditions will result in a tendency to equate price and cost. In our present example, since the only cost of production (ignoring the entrepreneurs' remuneration) is that incurred for labour, since we have (tacitly) assumed that the labour of all workers is of the same quality, and since we have (again tacitly) assumed that the technical coefficients are the same for all entrepreneurs, we find that price must tend to be equal to labour cost.

Price and cost. What we have still to examine is the effect of the (temporary) inequalities between price and cost which have been mentioned. We supposed, at one point in our argument, that price and cost were not equal. But this did not mean that when the markets for product and factor were viewed, each on its own, equilibrium did not exist. In each market

a certain demand met a certain supply, and a price was formed. But the fact that the two markets were related to one another meant that each price had to satisfy not only the supply and demand conditions in its own market but also those in the other. The prices could only be stable if they were in equilibrium relative to both markets. This mutual equilibrium is brought about by the operation of the law of cost. In one set of circumstances the entrepreneurs suffer losses, and they therefore change their demand in the one, and their supply in the other market. Under different conditions they make large gains; this will also lead to a change in the demand in one and in the supply in the other market. In either case the tendency is for readjustment to take place until cost and price are equated.

What, however, is the position while this readjustment is taking place, and before it is completed? In other words, why is it that inequalities can occur? The reason for the absence of equilibrium between the two markets (it is this absence which brings the law of cost into operation) is the fact that the entrepreneur will arrange his activities according to his estimates of the price-cost relationship. It may well be, therefore, that his estimate proves wrong, i.e. that he may have incurred costs which the price he can obtain will not cover.

True enough, any such error will tend to be corrected. When the entrepreneur comes to repeat the process of production he will, wiser for his experience,

act in the way described in the law of cost: he will offer less in wages and he will reduce his supply of the product in order to raise the price. This constant revision of the terms on which he is willing to buy and sell, based on a comparison of the entrepreneur's estimates with reality, is, in fact, what is meant by the mechanism of the law of cost. While, therefore, prices and costs tend to equality, divergence between them is possible: for no one will pay an entrepreneur a particular price simply because he has incurred certain costs.

This analysis enables us now more precisely to answer a common question: Does cost determine price, or not? Our answer is that costs and prices are interdependent and tend to equality. Costs will determine price only within the limits of the evaluations of the buyers of the product from which the objective scale results. Beyond that limit, costs do not determine price. Price thus forms an upper limit to costs, but costs do not form a lower limit to price. The equalization of the two is brought about by entrepreneurs producing according to estimates formed on the basis of existing prices and costs. Should that estimate turn out to be wrong, readjustment can only be made by a change in the supply of the product: when costs fall in relation to price, supply will increase and price will tend to fall; when costs rise in relation to price, supply will diminish and price will tend to rise.

An error once made is irrevocable. The answer to

the question whether cost determines price depends, therefore, on what we understand by cost. As Wicksteed has pointed out, if we mean by cost of production that expenditure which has already been incurred, we must conclude that it does not affect price. The price will be determined by the place which the product occupies on the objective scale. But cost, in the sense of the ratio between the price of the product and the price of anything which can be produced with the same outlay on factors, will determine whether the entrepreneur will produce that particular good or not and thus influence price. In the long run, therefore, both cost and price express, in different ways, the relative value of different goods. //

Opportunity cost. The relation just outlined is known as the *opportunity cost principle*. Its significance becomes clearer when we consider the production of more than one good by the same factor. Suppose, for example, that the entrepreneurs whom we have been considering had had the choice of supplying window-cleaning or 'charing'. In this case it will be the relative state of the two markets which will determine how much of each should be produced and what the prices will be, in exactly the same way in which, as we saw in section 2 of this chapter, the isolated economy chooses between two goods produced by the same factor.

If our entrepreneur discovers, after a time, that his estimates of the window-cleaning market were unduly optimistic, he may consider the possibility of using the

productive factor, labour, to supply 'chars' instead of going on with the supply of window-cleaning; but the loss made by the past error is incurred once and for all and cannot affect the price which he can get.

The following example, which is a slight adaptation of one given by Professor Bilimovič, may serve to illustrate the point in a more general way. Suppose there are three producers, A_1 , A_2 , and A_3 , who possess respectively 6, 6, and 4 units of a factor L . This factor can be used to produce either X or Y ; and the technical coefficients are the same for both, viz. 1 X requires 1 L and 1 Y also requires 1 L . We assume A_1 and A_2 to produce X and A_3 to produce Y . The total output will then be 12 X and 4 Y . Now suppose that the prices are: 1 unit of money buys 2 X and 1 unit of money buys 1 Y ; in other words, the ratio of exchange is 2 X : 1 Y . At this ratio (provided the objective scale is relevant to all three producers, i.e. providing that they are consumers of both X and Y), we can represent the scales of preferences as follows: $(X_1 + X_2) > Y_1 > (X_3 + X_4) > Y_2 > \dots$. A_1 and A_2 will then consume 4 X each and offer 2 X each in exchange for 1 Y . A_3 , on the other hand, will consume 2 Y and exchange the other 2 Y for 4 X . Supply and demand are equated and, taken in isolation, the markets for X and Y are in equilibrium.

But they are not in stable equilibrium, since the same factor produces different values in the two lines of production. In other words, the producer of Y fares relatively better than the producers of X . A change

will therefore take place in the relative amounts produced. A_1 and A_2 will enter the production of Y up to that point at which the same factor L produces the same value in the production of both goods. That will be the case when A_1 and A_2 produce $4X$ and $2Y$ each and A_3 produces $4Y$. The total output will then be $8X$ and $8Y$; the price will become $1 : 1$; A_1 and A_2 will, between them, demand $2Y$ in exchange for $2X$, A_3 will demand $2X$ and offer $2Y$. Demand and supply are equated and the two markets are in equilibrium. This time, however, it is an equilibrium which can last, for the economic optimum is reached in accordance with the second law of Gossen: the marginal utility produced by the factor in each line of production is the same as that in every other.

We have, in our last example, assumed that the technical coefficients were the same in the production of both commodities. Little is altered if we assume different technical coefficients for each good. It is unnecessary to give another example, for the analysis given will apply, provided that the technical coefficients are the same for each producer.

5. PRODUCER'S SURPLUS

Differences in technical coefficients. We must now turn to the case in which this last proviso is not fulfilled. That is to say, we drop the assumption that each producer has the same technical coefficients. From this it will follow that, where the stocks of the factors are fixed for each producer, a change over

from one line of production into another may be impossible. We speak in that case of an *incomplete mobility* of the factor. It will be shown that this incomplete mobility may be present for other reasons as well.

In order to study the implications of the existence of different technical coefficients (while still maintaining the condition that each good is produced by one factor only) we can either assume that each producer has a different factor, each of which is capable of producing the same good, though with different coefficients; or we can assume that each producer owns units of the same factor, but that these units are of differing qualities: the factor is heterogeneous. The former assumption is the simpler for purposes of exposition.

The first effect of the differences of coefficients is that we have now to deal with a multiplicity of cost levels. The costs of production of the same good (if we assume that the price of the factors is the same—as we should have to if we are postulating a competitive factor market—then costs of production will mean actual money outlays) will be different for each producer.

How, in this case, will the law of cost operate? The statement that the price will be equal to cost can no longer be regarded as sufficient, since there are different costs. When we say that there are producers with different costs, we imply, of course, that they all produce for the market, i.e. that they can all sell their

products. For each producer the law of cost will hold good. Each one will have to get a price which will cover cost, otherwise he will have to reduce his output or cease production altogether.

Marginal cost. This is true of all producers regardless of their costs. It is true of those with the highest, as well as of those with the lowest costs. While, however, costs may differ, we know that in the market there can only be one price for the same good. What will that price have to be in order to fulfil the condition that the law of cost must hold good for all producers? It clearly must be the price which is equal to the cost of the weakest producer, i.e. to the one who produces with the highest cost.

It seems paradoxical, at first sight, to say that price tends to equal the highest cost of production, but a little reflection will make the paradox disappear. For it is the highest cost of production, given a certain demand, to which price will be equal. In other words, the demand cannot be met by the supply of the lowest-cost producer alone; other portions of the supply must be resorted to. These other portions of supply can only be made available by having recourse to sources of supply which carry a progressively higher cost. And the cost of production of the marginal unit of the supply—the *marginal cost of production*—will determine the price.

It should be clear that the exact point at which the margin will lie at any particular time will vary. If, for some reason, such as a change in consumer's

preferences, the demand for the product declines sharply, price will fall. This fall will bring the law of cost into operation for those producers who are nearest the margin, for in their case the fall in price will produce losses. They will tend to go out of business, and the diminished demand will now be met by a smaller number of suppliers. But they, too, will continue to produce with different costs; and again a highest-cost producer will be present. There will still be a marginal cost of production, although, of course, the margin is now at a different level. Alternatively, an increase in demand will raise price and thereby encourage producers who were previously extra-marginal to enter this line of production. In either case, we find that the law of cost still operates; but it does so at the margin.

The existence of different costs of production and the working of the law of cost at the margin means, however, that different producers will achieve different remunerations. Those at the margin will only achieve that remuneration which makes it just worth their while to continue to produce. Those who produce with the lowest cost will obtain a high differential gain; this will get smaller as we near the marginal producer. The situation is somewhat analogous to that outlined in the previous chapter when seven buyers and seven sellers, of different strengths, were assumed to be bargaining. But while we found that we could not then speak in any quantitatively precise sense of consumer's surplus we can speak of a *producer's*

surplus in our present case. Here it is clearly a measurable difference between money outlays and money receipts.

Comparative costs. A further conclusion which we can draw in this example is that where technical coefficients differ for each producer of the same good, the same factor will produce different values for each producer. And if there are more than one good which can be produced by the same factor while the technical coefficients are still different for each good and for each producer, the previous analysis which showed that the law of cost will continue to operate at the margin will apply.

But what happens to our previously discovered tendency for the factor to be distributed among the different uses in such a way that it produces the same value in both? Let us suppose that A_1 , A_2 , and A_3 possess, respectively, 40, 25, and 28 units of a factor L which can produce either X or Y . The technical coefficients are different for each producer and for each good.

Let A_1	have technical coefficients	$\left\{ \begin{array}{l} 4L \text{ for } 1X \\ 8L \text{ for } 1Y \end{array} \right.$
A_2	„ „	$\left\{ \begin{array}{l} 5L \text{ for } 1X \\ 6L \text{ for } 1Y \end{array} \right.$
A_3	„ „	$\left\{ \begin{array}{l} 3L \text{ for } 1X \\ 2L \text{ for } 1Y \end{array} \right.$

If the price is such that $1X$ exchanges for $1Y$, production will be as follows: A_1 will produce $10X$; A_2 will produce $5X$, A_3 will produce $14Y$. With this output

—assuming again that the objective scale is relevant to all— A_1 will demand 5Y and offer 5X; A_2 will demand 2Y and offer 2X, and A_3 will demand 7X and offer 7Y. Supply and demand are thus equated and the price will be in equilibrium.

But with this output, L produces different values not only for each producer, but also in each line of production. In regard to the first difference, we have already seen that, with varying technical coefficients, it is inevitable that each producer who produces at a less cost than the marginal producer, will get a surplus. But since in our example the factor L produces $\frac{1}{2}$ a value unit in the production of Y, but only $\frac{1}{4}$ or $\frac{1}{5}$ in that of X, why does it not move from the production of X into the production of Y until it produces the same value in each?

The answer is simple again: the differences in the technical coefficients together with the fixity of stocks of factors. Each producer will still endeavour to produce in such a way as to get the best value out of the factors which he possesses. In a previous example in which technical coefficients were equal for each producer of the same good, this was achieved by using the factor for the production of that which, with existing prices, gave its owner, or user, the highest value. This is still true in our present example with the qualification that each producer will now use the factor for the production of that good which he can produce at a *lower comparative cost* than other producers. He will then rely on acquiring by exchange that good

which he himself can only produce at a higher comparative cost than others. In the case we have studied, A_3 had a lower cost of production than A_1 and A_2 for both X and Y. But he did not, on that account, produce them both and remain self-sufficient. He obtained a greater advantage by concentrating on the production of Y and getting X from producers who could only produce it at a higher cost than he might have done.

The principle of comparative cost is of great importance in explaining what are, at first sight, the paradoxes which one meets in business life. Whenever we wonder why people should make certain goods, sell them, and then buy others, we should ask ourselves whether, balancing the different costs against those of other producers (expressed in prices), it is not better to relinquish the production of that which one can make more cheaply than others for that which one can make very much more cheaply than others.

One woman may be able to produce the same frock more quickly than another woman. But if she could produce even more quickly a design for a fashion journal, and if the prices of the two goods are equal, then she would obviously be better off in producing the design. In other words, while she would get a surplus over the marginal producer if she made a frock, she would get an even greater surplus over the marginal producer if she produced a design.

International trade. The theory of comparative

costs is of special importance in explaining international trade. While this is a special subject which cannot be treated in detail in an exposition of the elementary principles of economics, the relevance of the above discussion to exchange between two countries may be pointed out.

The reasons for exchange between two countries are no different from those for exchange in general. They are different evaluations of the same good; different stocks of factors; different technical coefficients and, therefore, different costs. Suppose, for example, that countries A and B both possess a factor L and that the factor could produce X and Y. According to the law of cost, under competitive conditions the prices of the two commodities would, when costs vary for each producer, be uniform and would be proportional to the marginal costs of production. If labour could move freely from one producer to another it would flow to that producer who could produce most cheaply and he would supply the total demand; and if it could flow freely from the production of one product into that of another, it would be so distributed as to produce the same value, at the margin, in each.

But in international trade the assumption of perfect mobility of the factor from one producer to another does not hold. In particular, if that factor is labour, there are many barriers, of language and custom as well as those artificially created, which impede the free flow of labour according to the competition of the producers for its services. We must therefore assume

determine production. This would be fallacious. Technical coefficients are only important because, in the shape of costs, they enter into the process of choice of the producer. But they are by no means the only factors to be taken into account in that choice. In fact, it has been said that while technical data determine what alternative opportunities there are, the choice between these opportunities is an economic problem.

At this point a special case arises: one in which the range of choice which technique offers is small or non-existent. The problems of such a case are not, of course, different in kind from those studied in the preceding section; but the complete absence of choice, or its very narrow range, make them worthy of special attention.

The case referred to is that of *joint products*. There are many instances of goods which can only be produced jointly. The production of meat carries with it the production of hides; the production of mutton, that of wool; the production of coal-gas, that of coke and tar.

Fixed proportions. Let us first examine the case where no choice is open; the two goods are produced together in unvarying proportions from the same factor. (We continue to assume, for the sake of simplicity, that only one factor is used.) Examples of this kind are difficult to find; that of cotton-seed and cotton has long been regarded as the outstanding one. Fixed proportions imply that to produce them together involves exactly the same cost as to produce

each one singly. In a case of this kind it is impossible to assign costs to each separate product, i.e. to have separate reserve prices of the producers for each product. The total cost, at the margin, will influence the total of prices; but the proportion which the price of each bears to the total will be determined by demand conditions alone. If, for example, the production of 10 units of A always carries with it the production of 5 units of B and the marginal cost of production is 5 for 1 unit of A plus $\frac{1}{2}$ unit of B, all we can say, according to the law of cost, is that the combined prices will have to be such as to cover the marginal cost of 5. The prices may be made up of 4 for 1 unit of A and 1 for $\frac{1}{2}$ unit of B, or in any other combination. There is nothing in the law of cost which will determine the combination.

But we can explain the proportions by the fact that the prices for each product will be determined not only by the separate demand conditions, but will also be connected, since the products are linked on the supply side. A decline in the demand for one product will lower its price, and this will tend to reduce its supply. The supply of the other product would also be reduced, and if the demand for it remains the same, the price will tend to rise. Conversely, a rise in the demand for A by stimulating, through a rise in price, its supply will also stimulate the supply of B, the price of which, in the absence of any change in the demand for it, will fall.

We can reformulate the above statement in this

way: If the demand for one of two joint products falls, the supply of both will be decreased, unless the demand for the other rises sufficiently to enable its price to carry a greater share of the cost, and vice versa. It will depend on the respective elasticities of demand in each particular case, what exactly the changes will be in the relative prices and, therefore, in the relative share of the total cost which each product bears.

The case of two goods produced jointly in fixed proportions is rightly regarded as an exceptional one, but the general rule laid down above will hold in a greater or less degree of all joint products. According to the degree of variability of the proportions, an increase (or decrease) in the demand for one will tend to bring about a more or less proportionate increase (or decrease) in the supply of the other.

Variable proportions. But when the proportions can be altered, the cost of producing both is no longer that of producing each one singly. In other words it is possible to discover the changes in the total cost of production arising from changes in the amounts of each good produced. But once we can do that, we can find the marginal cost of each good, which will be the diminution in the total cost occasioned, in Marshall's words, 'by so modifying the proportions as slightly to diminish the amount of one . . . without affecting the amounts of the others'.

Suppose, for example, that per unit cost one can obtain either $4A+2B$ or $5A+1B$. Then a total cost of 5 could produce $20A+10B$, and a total cost of

4 could produce $20A + 4B$. In other words, to produce an extra 6 of B without altering the amount of A produced would cost 1; the marginal cost of B is, therefore, $\frac{1}{6}$. It can be found for A in the same way; and when we have found the two marginal costs, the prices (dependent on the relative demands for A and B) will determine the proportions in which A and B will, in fact, be produced. The only modification which is necessary is to introduce the costs which are peculiar to each product. In the case of mutton and wool, for example, there are, in addition to the joint cost of breeding the sheep, the separate costs of killing the animal and preparing the carcass and those of shearing, &c. The price of each product will have to cover its special cost plus its appropriate share of the joint cost.

Joint products are very important in practice; and their costing presents one of the most difficult problems of business. The demands for them vary greatly from time to time. What is at one time the by-product which is difficult to dispose of, may become the main product. The invention of refrigeration made it possible to transport mutton from New Zealand (where previously it had been a by-product of wool-production), and the increase in the demand for mutton led to a substitution of cross-bred sheep, which yield more meat, for Merino sheep which yield more wool. Cotton lint, which is a by-product of the seed-crushing industry, might go up in demand (e.g. in time of war) and the position of seed-crushing would be greatly

altered. The supply of seed would go up and, unless the demand for it rose, its price would fall. As far as the oil and cake made from the seed are concerned, it would depend partly on the elasticity of the demand for it, and partly on the size of its special costs of production in relation to its share of the joint cost, whether its supply would go up. If the demand is elastic, and if its special costs are not very high in relation to its share of the joint cost, supply may expand, price will fall, and demand will be stimulated. If, however, demand is inelastic and the special costs are proportionately very high, then even a very high price for lint will not make it worth while to incur the additional special costs involved in an increased supply of seed-cake and oil.

The importance of elasticity of demand under conditions of joint cost (when the proportions of output are more or less invariable) is seen particularly clearly in the case of a monopolist's price policy. Here, joint production and differing elasticities of demand for the joint products may lead to a policy of charging prices in accordance with demand conditions without endeavouring to ascertain the proportions of total cost (though not special costs), provided that the sum total of prices covers the sum total of costs. The outstanding example of this policy is that of transport charges, where the principle of 'what the traffic will bear' operates. The transport of different types of goods is not charged for in accordance with their respective shares of the total cost of transport, but in

accordance with the respective demands. Where the prices for the transport of certain goods are fixed very low by legislation the principle of 'what the traffic will bear' will lead to a very high price for transport in those cases in which demand is highly inelastic.

7. COMPLEX PRODUCTION

Throughout our discussion of production so far, we have assumed that only one factor was used. It was pointed out that while this was clearly the simplest and the most fundamental case, it was one which was either unlikely to be found in reality at all (our window-cleaning example), or, if it was, was not to be regarded as typical of the present organization of industry. We can take it as the general rule that modern production is production by means of a combination of factors. What we have now to examine is the extent to which our conclusions can be applied to this present case. The law of cost has shown the connexion between product and factor prices. Is it relevant and, if so, with what modifications to complex production?

Variable combinations of factors. There was, of course, one point in our previous analysis at which the use of a combination of factors was assumed: in the case of an isolated economy using land and labour to produce goods. But that example did not raise any special problems because we assumed that the proportions of the productive combination could not be varied. We must analyse the more general case when

there is variability of the combination, and our results will then be equally applicable to the isolated and to the exchange economy. It seems clear that the law of cost will be applicable in a general way. Once again, the entrepreneur will transform a demand for consumption goods into a demand for factors, and a supply of factors into a supply of consumption goods; and, in so doing, he will arrange his demand and supply so as to fulfil the conditions of the law of cost. The difference that arises in this case is, however, this: granted that total cost, at the margin, must equal price, what share will each factor have of that total? Or, to put essentially the same question in a different form, in what particular quantities will the factors be used, since the possibilities of combination are many: much of one factor may be used for little of another; one factor may be entirely replaced by another.

The law of diminishing returns. To answer this question we have to look a little more closely at the technical coefficients. In our previous examples we tacitly assumed that the technical coefficients remained unchanged for all sizes of output. The general assumption was one of *constant returns* (i.e. constant physical output per unit of factor) or *constant costs* (constant money outlay per unit of product).

Is this assumption still justified when a variable combination of factors is used? The answer to this question was first given in relation to agricultural production and was formulated as the *law of diminishing returns* or *increasing costs*. In the more general form

which is common to-day it becomes one of the most fundamental laws of economics.

First systematically developed by Turgot, and introduced into this country by West, the law states that beyond a certain point an increased application of labour to the same piece of land yields a less than proportionately increased return. Thus if 10 labourers could raise 20 bushels of wheat per acre, 12 labourers might raise 23, 15 might produce 27, and 20 might succeed in raising 32 bushels. While every additional unit of labour added something to the product, after a certain point it did so only in a diminishing ratio, thus leading to a decline in the size of the *average product* per unit of factor.

Attempts have sometimes been made to prove or support this law by reference to certain technical facts concerning agricultural production. But the law of diminishing returns is not dependent on such proof. It follows from the simple observation that land does command a price, that farmers are often anxious to obtain more land (this shows itself in land-hunger), and that no farmer uses only the best portions of his land, but cultivates poorer portions as well. If the law of diminishing returns did not hold good, the return from any portion of land could be increased indefinitely by an increase only of the amount of labour used on it. But we know that this is not so. We know that after a certain point, if an increase in output is desired, it will be profitable to take more land into cultivation, even if it is poorer land.

What is true of the returns under the above conditions, when the amount of labour was varied while the amount of land was fixed, is equally true in the converse case, that of an increase of land without an increase in labour. A doubling of the amount of land to be cultivated by the same amount of labour will certainly not produce double the output. It is therefore possible to widen the law of diminishing returns, and to make it applicable to all production by means of a combination of factors. We conclude that if in a combination of factors one factor remains unchanged while the others are increased, additional units of those factors will, after a certain point, yield a smaller return than preceding ones.

Increasing returns. Serious objections have been raised against this law on the ground that experience has shown many cases in which an increase in production has been followed not by a rise but by a fall in cost per unit, thus revealing a condition of *increasing returns*, or *decreasing costs*.

At one time it was even believed that there were, in fact, these two laws, operating side by side, the one applicable to agriculture, the other to manufacturing industry. In agriculture, it was said, diminishing returns operate, while in industry increasing returns are to be found. Later experience suggested some modifications of this view. The improvements in the technique of farming led to an increased output without any increase in the area under cultivation. And while 'mass-production' methods in industry seemed

to provide striking examples of increasing returns, it was found that there were occasions, even in manufacture, when an increase in output required not only an increase of labour, but an addition to the factory buildings, an increase in the amount of raw materials, &c., if the returns were to be at least proportionate. It was therefore thought that in both agriculture and industry increasing and diminishing returns might operate. The point at which the former would turn into the latter might be reached sooner in agriculture, but it would exist in industry as well.

The problem with which these laws were designed to grapple was the relation between cost and output. The supply-price relationship, which we studied in the theory of exchange (when variability of stocks was excluded), was found to be insufficient to explain the changes of supply in response to changes in demand and price and, in return, the effects of such changes on price and demand in the conditions of complex production. Not only were these conditions more involved owing to the possibility of altering the combinations of factors which were employed, but they were also complicated by the fact that there were numbers of independent producers in each line of production, thus leading to the further distinction between the cost-output relations of the industry and those of the individual firm.

As regards the industry, until recently, the main change from the classical position was a generalization of both laws. From being first applied exclusively to

land, the law of diminishing returns was said to hold, as was pointed out above, in all cases in which a constant factor was employed.

Indivisible Factors. Increasing returns were regarded as part of the principle of division of labour. It was recognized, in the first place, that diminishing returns began only 'beyond a certain point'. Just as in Gossen's first law we found that there was a *minimum sensibile* below which the law did not operate, so also in the case of production there must be a minimum combination of factors which will be of significance. It is only beyond that minimum that diminishing returns operate. The utility of motor-cars will not begin to decline until the want for at least one whole car (and not a minute fraction of one) has been satisfied. Similarly, to produce a bushel of wheat will require a minimum amount of land and labour. These provisos rest on the *incomplete divisibility* of wants and factors. In the second place, it was made clear that the law of diminishing returns is limited by the qualification, 'other things being equal', in exactly the same way as all the laws we have discussed, in particular the law of diminishing utility.

We could, then, say that, from a position of equilibrium, an increased supply of a good can only be brought about by withdrawing factors of production from other uses and thereby increasing cost. If, on the other hand, there is an increase in the demand for one good, accompanied by a fall in the demand for other goods (i.e. other things do not remain equal),

then more factors will be available to supply the increased demand, and a new productive combination is possible. It is quite conceivable that the technical facts of this new combination will be such as to make it more productive than the old one; in this case, it might be possible to speak of cost having decreased, or returns increased. But it must be clearly understood that what we are doing is to compare cost in one situation and cost in an entirely changed situation. The cost of production per unit of one good will be different for different sizes of output, if the output of other goods is also different. It does not mean that the law of diminishing returns is thereby invalidated. Even where cost has fallen (because all the circumstances have changed) a further increase in output, without the change in the accompanying circumstances continuing, can only be brought about under conditions of diminishing returns.

The variations just examined which gave rise to decreasing cost lie on the demand side. But increasing returns much more often refer to variations in the technical conditions of production—what are generally called the advantages of *large-scale production*. It is a technical fact that extension of division of labour increases productivity. It is equally a fact that the extent to which division of labour can be carried on depends upon the market which it serves. As Adam Smith pointed out, the wider the market, the greater the division of labour. An extension of the market makes it possible to introduce

greater division of labour and thereby to employ the same factors more successfully, i.e. to produce at a lower cost.

It is here, as has more recently been pointed out, that the incomplete divisibility of factors noted above becomes of importance. Because of it, the adoption of a greater division of labour must wait until the market has extended to a certain size; or, in other words, the increase in the division of labour cannot take place in infinitely small stages, but only in 'jumps'. Once a higher stage in the division of labour has been reached, any increase in output resulting from an extension of the market can be met at a decreasing cost per unit.

A building erected to accommodate the production of 50 units of a good may be big enough for an output twice as great. Therefore, any increase in output between 50 and 100 will reduce the share of the cost of the building which each unit of output has to bear. Beyond an output of 100, costs per unit will again begin to rise; since the building will no longer be big enough, there will be overcrowding with a consequent diminution in efficiency. To cope with it, it may be necessary to put up another building, again with a capacity of 100; and until that capacity is reached cost per unit will again decline.

Increasing returns are thus due to discontinuities in the introduction of division of labour which result from the indivisibility of the factors of production. In practice this indivisibility becomes more marked

when we consider (at a later stage) capitalistic production.

Individual increasing returns. One difficulty soon became apparent: the existence of increasing returns in the individual firm (due to the internal economies of large-scale production) was found to be incompatible with conditions of perfect competition. For if, by expanding output, the individual producer could lower his cost, he might take active steps to increase his individual market—by lowering price, for example. And it would pay him to do so, as long as what he sacrificed in order to gain his increased market was less than the reduction in cost which he could thereby effect.

Suppose, for example, that there had occurred a sudden fall in demand. The firms in which increasing returns are operating will find that with every diminution in output, the cost per unit would go up and their losses would increase. The only way to bring about a reduction in cost is to increase output. But every increase in output will tend to depress prices still further. The only way out of this difficulty seems to be some measure by which a reduction in total supply can be brought about, such that the rise in price more than offsets the rise in cost, even though for each individual entrepreneur this would involve losses.

Cut-throat competition. The measure for achieving this is a restriction of competition among entrepreneurs. No individual entrepreneur can hope to

restrict production in order to raise prices. Others would at once undercut him, for if they are not reducing their output, they are producing at lower cost. In fact, under competitive conditions, the existence of increasing returns in the firm must lead to attempts at undercutting. Every entrepreneur will try, by lowering his price, to attract so large a part of the demand to himself that he will not have to restrict his output and thereby increase his cost per unit. This situation is known as *suicidal*, or *cut-throat competition*, since there is no reason why this price-cutting should not be resorted to by all producers, thus leading to continually increasing losses all round. The ultimate result will be a disappearance of all those producers who are unable to stand the continued losses, with a consequent diminution of the total supply. One producer will be left and will now have an opportunity to restrict production and raise prices again. Or, and this is more likely in practice, before all producers but the strongest have been driven out, an attempt will be made to form a *combination* of producers which will be able to control the market as a monopolist. One of the important results, therefore, of the existence of increasing returns for the firm is the tendency towards the restriction of competition whenever changes in demand produce a price-cost ratio which involves losses. This will be all the more so if the demand for the particular commodity is elastic, i.e. if a reduction in price will increase the demand more than in proportion. It is likely, in that case, that the

best policy for the individual firm will be a large supply at a comparatively low price.

What circumstances are likely to make demand elastic? The answer usually given, which turns on the difference between necessities and luxuries, we have found to be unsatisfactory. A better division is that between goods for which substitutes, which are not markedly inferior, are, or are not, readily available. The former will have a more elastic demand than the latter; and the general rules about the relations of the demand for substitutes (including those which refer to changes in total income) will have to be applied before we can say, in any individual instance, how quickly demand will react to price changes, or vice versa. Different grades of the same good are some of the most likely substitutes for one another. But any circumstance which, in the minds of buyers, marks off one good from another will diminish the possibility of substitution and will reduce the elasticity of the demand for one as compared with the demand for the other. Branded and widely advertised goods are striking examples. The effect of advertising a particular brand is, as it were, to insulate it and to make the demand for it of less elasticity than infinity in the critical range (as pointed out on p. 105). The advertisement then becomes something in the nature of a joint product; and the price of the goods will have to be high enough to bear the total joint costs of production. From the seller's point of view, advertisement has the purpose of preventing substitution—an aim

which might conceivably be achieved by engaging in price competition. The consequences of the seller's choice of methods are far-reaching, but their problems are too complex to be discussed here.

Diminishing and increasing return industries. It is, at any rate, clear that increasing returns, due to internal economies, and conditions of perfect competition can never exist together. Stress has therefore been laid—since Marshall—upon 'external' economies, that is to say, the economies derived by all individual firms in an industry from the growth of that industry as a whole. This does not mean (as it is sometimes understood) general industrial progress, but the kind of progress which, while external to the individual firm, affects the particular industry.

If diminishing and increasing returns are understood in the senses given above, it becomes much more difficult to divide lines of production according to whether they show the one or the other law in operation. If we base our definition of an industry upon a particular factor which the industry uses (such as agriculture, or the extractive industries), diminishing returns are more likely to be in evidence; while increasing returns are more probable when we define an industry according to a particular kind of finished consumption good.

Indeed, Mr. Sraffa has recently argued most forcibly that diminishing returns apply only to that small 'class of commodities in the production of which the

whole of a factor of production is employed'. He has shown that, in all other cases, there is either a negligible change of cost following upon a small change in production; or there will be a repercussive effect upon the cost and price of, and therefore demand for, other industries using the same constant factor: other things do not remain equal. Similarly, with increasing returns, only the rare cases are left of economies which are internal to the industry but external to the firm. For the conditions of perfect competition, then, it would seem that neither law is adequate.

Thus, in any further discussion of the relation between price, cost, and output, it is necessary to consider the actions of the individual firm under conditions of both perfect and imperfect competition.

Price, cost, and output in perfect competition. In perfect competition no individual entrepreneur will be able to affect price. Therefore, whatever he succeeds in selling, he obtains for each additional unit sold the ruling market price. He will, therefore, expand his output up to the point at which what he adds to his receipts by selling the last unit (the price) is equal to what he adds to his total cost of production by producing the last unit (the *marginal cost*). In perfect competition, therefore, price will equal marginal cost of production.

But it is also clear that each firm will endeavour to obtain the maximum revenue: that is to say, it will endeavour to minimize its *average cost* of production. A firm will expand output so long as it

can thereby reduce its average cost; and it will stop expanding if it were thereby to increase its average cost of production. The point, therefore, at which it will stop will be the point at which the cost of an additional unit is neither greater nor less than the average cost: in other words, it will stop at that scale of output at which average and marginal cost are equal. Since, in perfect competition, price equals marginal cost, it follows that it must also equal average cost.

In perfect competition price must, as we have seen, also equal the average cost of production in the industry. If it did not, there would be profits or losses which, given perfect mobility, would lead to an influx or outflow of entrepreneurs.

Price, cost, and output in imperfect competition.

The main characteristic of imperfect competition is that the individual seller can affect the price. It will still be true that the firm will expand up to the point at which its marginal cost will equal the receipt for the marginal unit sold. But in imperfect competition this will no longer be the price, since, in order to sell more, the firm will have to lower the price. It will therefore stop at some point before price has fallen to marginal cost.

When can imperfect competition be said to exist? In the first place, of course, when there are few suppliers, each of whom controls a considerable portion of the supply. In the second place, there may be market imperfections, such as transport costs, lack of knowledge of the sources of supply on the part of

buyers, and differences in the products of individual suppliers, regardless of whether these are real or imaginary.

Even when there are many individual producers the probability, under modern conditions of production, is that each individual producer will be working with increasing returns. What he has to do in order to increase his sales is to widen his market, either by reducing price or by the expenditure of so-called 'sales-costs'. The degree of imperfection of the market, the chief cause of which is the attachment of buyers to the products of an individual producer, determines the extent to which each individual producer can expand his own sales. In perfect competition we have found that the elasticity of the demand for the product of an individual producer, within a certain range, is infinite. In complete monopoly the elasticity of demand would be unity within that range. In imperfect competition it is somewhere between these two. Each individual producer is, then, subject to some degree of competition, both from his own rivals as well as from producers of alternative goods on which the consumer might spend his money. The degree of attachment of his own clientèle will determine the extent to which his price may exceed those of the rivals in the same line of production. In addition, there will be a limit beyond which the price of all similar commodities cannot rise without driving buyers out of the whole market altogether.

Mr. Sraffa has demonstrated that, in the conditions of imperfect competition which are most likely to arise

in practice (i.e. in which individual demand is attached), output will tend to be lower and price higher than in conditions of perfect competition. If each producer raises his price, buyers, finding substitutes rising, will continue to buy the product they have hitherto preferred. There will thus tend to be a progressive all-round increase in price up to the point at which total demand begins to diminish. For fear of retaliation, no firm is likely to begin to cut price unless demand is very elastic and costs are rapidly decreasing. In these conditions price-cutting will tend to occur and, as we have seen, complete monopolization is likely to result. Where, therefore, increasing returns in the firm are very marked, we can say that there will be a tendency to monopolization or to the use of means (e.g. advertising) for insulating individual demands. In the latter case the price paid for stability is, as has been pointed out, lower output and higher prices with, in addition, the encouragement of variety, which may be largely illusory.

The principle of marginal productivity. We must now return to the question of the remuneration of the factors. The last increment of a factor used will be the marginal one; its addition to the total product will be its *marginal product*. The marginal product of a factor, in other words, will be the total output of a productive combination in which that factor is used, minus the total output of a combination in which one unit less of that factor, but the same amount of other factors, are employed.

The introduction of the concept of the marginal product enables us at once to answer the questions set at the beginning of this section: in what proportions does the entrepreneur use the factors, and how are cost and price equated? The operation of the law of cost involved a consideration of an increase or decrease in the supply of products by the entrepreneur, and a corresponding increase or decrease in his demand for factors. This will still be so; but, in deciding whether to produce an additional unit of the good, the entrepreneur will now have to decide how much of this factor or of that he is to use. In making this decision he will have to take into account the marginal product—the addition made to the total output by an additional unit of a factor.

If the price of a factor is lower than the price of its marginal product, more units of that factor will be employed; while any factor, the price of whose marginal product is below its price, will be less used. Both these tendencies will continue for all factors until the price of the marginal product and the price of a unit of the factor are equal.

This statement is known as the law of marginal productivity, and it represents nothing else than an adaptation to the conditions of production of Gossen's laws.

Marginal productivity in perfect and imperfect competition. It will be clear, however, from what has already been said that, in this form, the law will only hold good if there is perfect competition. Each entre-

preneur will continue to employ a factor so long as the cost of additional units of the factor is below the addition to the entrepreneur's total receipts caused by increasing output. In perfect competition this addition to total receipts, caused by each additional unit of output, will be the price. Thus, in perfect competition, each factor will be paid a price equal to the value of its marginal product.

In imperfect competition, the entrepreneur will still continue to employ a factor so long as he adds to his receipts at least as much as the cost of the additional units of the factor employed. But we know that when competition is imperfect the entrepreneur will have to reduce price in order to sell more. He will therefore stop increasing output before price has fallen to marginal cost. This means that he will stop employing a factor at some point before the value of the marginal product has fallen to the level of the price of the factor. In other words, in imperfect competition, a factor will be paid less than the value of its marginal product.

8. RENT

Before we turn to capitalistic production it is necessary to devote a little more attention to the problem of the prices of the factors of production. What share of the total product do the factors which have contributed to its production obtain? We have already, in a general way, answered this question, which is named the problem of distribution, by saying that the price of each factor will tend to be equal to its marginal

product. The further analysis of this statement must be made for land in the *theory of rent*, and for labour in the *theory of wages*.

The theory of rent is one of the earliest achievements of economic science. It goes back to Ricardo, and very little has been added to it since his day. The problem which has to be solved in relation to land is that of the existence of differences in the remuneration of different units of the same factor. Land obtains a remuneration because it is an economic good: it is available in limited quantity in relation to the demand for it. Its owner can therefore get a payment for what Ricardo called 'the permanent and indestructible properties of the soil' which will be measured by the marginal product of that land. Rent, therefore—and this applies not only to land but also to other factors of production—is simply the obverse of marginal productivity. As a payment accruing to a particular class of individuals, it is the result of the existence of private property in land.

We will suppose that a community settles on land of a certain quality which is capable of yielding 24 bushels of wheat per acre. So long as the land is superabundant it will be a free good and will not command a price—wheat, however, will have a price which will tend to be equal to its cost of production. If we suppose the cost of tilling the land to be represented by £3 per acre, the cost of wheat will tend to be 2s. 6d. per bushel.

If the community increases beyond a certain point, an increase in the production of wheat on the same

land can only be brought about with diminishing returns. In order to obviate that, more land has to be brought into cultivation. Now if the amount of land of the first quality is limited, recourse will have to be had to poorer land on which the yield is, say, only 20 bushels per acre. This second quality land is a free good, it will not command a price. But wheat will continue to have a price; and this price will now tend to equal a higher cost of production, that on the poorer quality land. With a cost of labour per acre unchanged at £3 per acre, the cost of wheat will now be 3s. od. per bushel. The price of wheat will be uniform, regardless of the conditions under which it is produced and will tend to equal the marginal cost of production, 3s. od.

The producers on the first-quality land will obtain a differential surplus of 6d., due to the superiority of the quality of their land over the marginal land. This producer's surplus is called *rent*; it expresses the difference in the productivity of an acre of land over that of the *marginal*, or *no-rent* acre.

Rent of land is thus nothing less than the producer's surplus which we have already met when we discussed the existence of differences in technical coefficients between different producers, except that we have here assumed differences in the quality of different units of the factor. Originally worked out for agricultural land, the concept is applicable to any case in which different producers make the same thing with different costs of production in relation to price. These differ-

ences will arise whenever there are variations in costs of production and whenever the total demand cannot be met with the lowest cost alone. A rental element can also be found in the remuneration of all factors of production which exhibit differences in quality. This will become even clearer when we look at the remuneration of labour.

9. WAGES

Wages are the remuneration of labour. Once again it becomes necessary to distinguish between the everyday and the technical uses of this term. This time, however, we can make a clear distinction by using the words *explicit* and *implicit*. Explicit wages are wages as normally understood, i.e. the contractual payment made to hired labour—a certain reward, per day, per week, per month, or per unit of output contracted for in advance between the owner of the factor labour and the entrepreneur who hires it.

The wage system. The existence of wages in this sense is dependent upon the existence of a certain social system such as was outlined in Part I. The wages system implies the existence of a labouring class which is forced by the conditions in which production takes place (private property in the means of production, division of labour, co-operative production, the entrepreneur, the market) to sell its services in the labour market; it implies also freedom of contract for labour. When these conditions are present, we have both the necessity for a sale of labour and the possibility of its being purchased.

Wages have also been thought of in a wider sense as the share of the total product attributable to the factor labour, whether hired or not. There is, for example, the independent craftsman who combines certain materials with his own efforts and produces a good which he then sells in the market. Part of his earnings derive from the expenditure of labour; and he can ascertain what part that is by deducting from his yearly income the amount which he could have earned had he hired out his labour in the market. Thus while implicit wages are always present where labour is used in production, it is not possible to measure them unless a labour market exists, i.e. unless regular dealings are taking place in the commodity labour power which put it on the communal scale and give it a price—explicit wages.

Real wages and money wages. It is, then, with explicit wages that we are concerned; and it is only through a study of the way in which they are determined that we are enabled to make any generalizations about implicit wages also. There is a further distinction which must be made. When we speak of the price paid to the labourer, we naturally have in mind the monetary reward obtained. But we must not forget that in this, as in the case of every exchange, money is only a medium which facilitates the exchange of goods and services. It is customary, therefore, to distinguish between the *money wages* which the worker gets, and his *real wages*, which are the things the money wages will buy. Sometimes the worker will

receive his remuneration partly in money and partly in goods and services, though the revolt against the *truck* system, when wages were entirely in kind, has made this method much rarer than it used to be.

It is important to keep in mind the difference between real and money wages, particularly when we wish to compare the earnings of labour at one time or place and at another time or place. The same money wage will represent a much lower real wage when the prices of the goods the worker consumes are high than when these prices are low. It is not easy to make allowance for these changes in prices, and no perfect method has yet been devised, for the simple reason that it is not possible to assume that the goods consumed are the same for all workers or even for the same worker all the time. And, strictly speaking, it is only each individual worker who can say how much real wage he is at any time getting. But, for purposes of general comparison, systems of *Index Numbers* can be devised which give a sufficiently good indication of changes. The prices at a particular time of what are assumed to be the articles of consumption of the workers are regarded as 100 per cent., and changes in the price of that composite bundle of goods are expressed as percentage changes. Care must be taken in interpreting these changes: their validity depends on the continued relevance, for purposes of consumption, of the bundle chosen.

Two further words of caution are necessary: one concerns the extent to which wages can be used for

the purpose of judging the workers' remuneration, the other the reliance on wages for the purpose of ascertaining the entrepreneur's cost. On the first point, it should be borne in mind that not only must money wages be translated into real wages before the remuneration of the worker can be assessed, but a clear distinction must also be made between wages and *earnings*: a wage rate of £3 a week may represent very low earnings if the worker works only two days a week.

Wages and the cost of labour. The second point is equally important and often gives rise to serious misunderstanding. To what extent do high wages represent a high cost of labour? It is clear that the entrepreneur will try to buy his labour as cheaply as he can; but what will determine the cheapness or otherwise to him of labour will be not so much the wage he pays, but rather the relation between the wage he pays and what he gets for it. It does not, therefore, follow that a high wage is necessarily a high price for the labour an entrepreneur buys. The efficiency of the worker, his output, determine to the employer whether the wage paid is a high or a low cost of labour. This distinction is of very great importance in many practical issues of to-day when attempts are made to compare the cost of production (as far as labour is concerned) of different producers, particularly when these producers are situated in different countries. Moreover, it must always be borne in mind that the proportion which labour cost bears to the whole will vary greatly from one industry to another.

The theory of wages. There are two important questions which the theory of wages has to answer: how are wages in general determined in the market? and how can we explain the differences between the wages ruling in different occupations? To the first, the answer is given in the general law of price, as formulated for the factors of production. The answer to the second is to be found in the theory of net advantages and non-competing groups.

The demand for labour, as we know, comes from the entrepreneur. He wishes to buy labour for purposes of production and he will arrange his offers in accordance with the law of cost. This, under conditions of complex production, becomes the principle of marginal productivity; and the highest price per unit of a given supply of labour which the entrepreneur will be willing to pay will depend upon the marginal productivity of labour; i.e. the difference to the total product made by withdrawing the last unit of labour from a given combination of factors.

The supply of labour. As regards the supply of labour, we may look upon that either as the total number of workers willing to hire themselves out for wages or the length of time (e.g. number of hours) which each one is willing to work, or, again, the intensity of the work performed by each worker. It is possible to include all three and to speak of the supply of labour as the amount of physical energy forthcoming at a given price. We have noted that, in the case of labour, we must be prepared to meet certain

exceptions to the general laws of exchange; in other words, we cannot always assume that the relation between the supply of labour and wages will be such that to call forth an additional supply will involve the offer of a higher wage, while down to a certain level, a lowering of the wage will result in a reduced supply. Indeed, it has recently been argued (by Professor Douglas) that statistical evidence shows that there is a considerable range within which earnings and the supply of labour are indirectly related. This, in other words, would mean that there is a range within which the supply of effort in terms of money income is inelastic: a rise in earnings tends to be 'taken out' more in leisure than in material goods.

The level below or above which the direct relation between supply and price will not hold is given by the minimum standard of living which the worker wishes to maintain. Below the wage capable of maintaining that standard no diminution in supply will take place—indeed, the supply may even increase; above, no increase in supply will take place and it may even diminish. The classical economists gave expression to the former fact by saying that there was a minimum subsistence level which formed the reserve price of the seller of labour power.

Modern economics has made the term minimum subsistence level have a conventional rather than an absolute physical quality. It represents the minimum standard of existence, including, at any rate, the minor comforts of life, to which the working class has

become accustomed. In introducing this element we have introduced a strong element of habit and custom into the determination of wages.

It is difficult to speak of reserve prices as the lowest limits at which the workers will be prepared to hire themselves out. The more complex production becomes, the less alternative openings will there be for the labourer, the less will he be in a position to withhold his labour, however low the price; and a decline of wages even to the minimum of physical subsistence postulated by the classical economists is not impossible.

If there is any sense in which the minimum level of physical subsistence can be said to form a lower limit to wages, the marginal product of labour, in any given set of conditions, certainly forms an upper limit. Wages will lie between these. Anything that tends to raise the workers' reserve prices will raise the level of wages nearer to the value of the marginal product; anything that tends to lower them will bring wages nearer to the minimum of physical subsistence. In other words, within the range set by the two extremes, wages will vary largely in accordance with the relative bargaining strength of workers and employers. A monopolistic situation on the employers' side will make them able to influence the wage level in a downward direction; a monopolistic organization on the workers' side, such as is aimed at through the formation of trade unions, will increase the bargaining strength of the workers and raise the level of wages nearer to the value

of the marginal product. When we have strong organizations on both sides, it is impossible to say beforehand where the level will lie—it is a case of bilateral monopoly. We conclude, with Professor Pigou, that there is a range of indeterminacy of wages.

The marginal product of labour. There is, however, one point which requires further analysis: the marginal product of labour. We have seen that it only has precise meaning in connexion with a given productive combination. But what if that productive combination changes? What if the entrepreneur increases the amount of other factors used? Labour is a complementary good as far as complex production is concerned, at any rate when we move from an equilibrium position, for, from that point, the increase of one factor will increase output only with diminishing returns. Thus an increase in one factor must raise the marginal product of labour and thereby increase the price which the entrepreneur is willing to pay for labour. In other words, we find the situation to be the same as that of all complementary goods: the value of the marginal product of labour does not depend only on its own supply, but also on that of the other complements. Any increase in their quantity must raise the relative marginal productivity of labour. A full discussion of this point involves the theory of capital; and it must, therefore, be postponed for a moment.

Wage differences. The point to which we must turn now is a different one. We have, so far, been con-

cerned with the wage-level in general. This, however, in practice seems to be of very little importance, for we find very wide variations in the wages ruling in different industries. We should have expected that competition between workers would quickly achieve uniformity of wages. A high wage in one occupation would quickly attract workers away from lowly paid employments; and the consequent rise in the supply of labour in one occupation and fall in the other would equalize their prices.

Net advantages. Even if there existed perfect mobility of labour between different occupations, we could not find a uniform level of wages, for the simple reason that wages constitute only the monetary remuneration of the worker. It has been pointed out by Adam Smith that the power of attraction or repulsion of different industries must be conceived of in wider terms than wages alone. We have to compare all the advantages (including wages) which an occupation offers with all the advantages (including wages) of others. Occupations which are unpleasant, arduous, or dangerous; those which involve special training and a long period of apprenticeship; those with insecure tenure: all these would be regarded as less advantageous, and wages in these employments would have to be correspondingly higher. On the other hand, wages might be low in occupations which offered other such compensating advantages as security of tenure, higher social position, long holidays, &c.

Thus differences in wages would simply be the

reflection of differences in the attractiveness of certain occupations. The tendency would be for competition among workers to equalize not money wages but *net advantages*. In addition, differences in money wages may be due to differences in the quality of different types of labour. When the supply of the best quality is fixed and recourse has to be had to inferior qualities, the superior grades will obtain a surplus remuneration which is in the nature of rent.

Non-competing groups. But equalization even of net advantages can hardly be said to be achieved in practice. Prestige and security of tenure often go together with a comparatively high wage, while some of the most arduous and insecure work is often the most lowly paid. The explanation of this phenomenon lies in the fact that labour is not perfectly mobile. As regards different countries and even different areas of the same country, we have seen that there are strong barriers in language, habit, and cost of transport which prevent the free flow of labour.

Perfect mobility is also absent when we consider different occupations. Entry into different employments is seldom free to all. Sometimes special skill is necessary which can be acquired only after a costly training; sometimes social position and background create an effective barrier to new-comers; sometimes there are artificial restrictions which may be imposed by certain industries themselves or may even be backed by legislation. The skill required for a craft involves a long apprenticeship, and only those mem-

bers of the working class with comparatively large incomes will be able to offer their children the opportunity of acquiring that skill. Again, the education which leads to a career in the professions is so expensive that, in spite of the increased aid available from public sources, it is not often that a child of any but a middle-class family can afford it. And for some professions, such as the Diplomatic Service, even a particular education is not enough without the backing of a certain social position.

Nor has entry to different occupations ever been left entirely free from deliberate regulation. It is true that the restrictions imposed by the craft guilds of the Middle Ages were largely broken down in the eighteenth and nineteenth centuries, but many of the professions are carrying on their tradition and, by imposing entrance examinations or demanding premiums from apprentices, are endeavouring to limit the entry of new-comers. Of late the actions of the State have also tended to diminish occupational mobility. A system of licensing, or even a complete bar to fresh entrants, is by no means uncommon to-day; although it generally affects the independent entrepreneur rather than the wage-earner. It is, nevertheless, of importance since it narrows still further the range of alternatives open to the worker.

10. CAPITALISTIC PRODUCTION

Production by roundabout methods. We are now in a position to embark upon the last part of our

analysis of production, the effect of the use of what have been called capitalistic methods. The explanation of a number of points has had to be left incomplete until an analysis of capital has been given. In the first place, in our discussion of the remuneration of land and of labour we have always assumed that these two factors are used directly to produce want-satisfying services; we have left out of account the possibility of indirect production mentioned in Part I. Secondly, we have found that increasing returns arose from the indivisibility of factors and, as will be shown presently, this is more likely to arise when production is indirect. Lastly, we have noticed that the marginal product of labour was dependent also on the quantity of other factors used in production; in the present section, we shall find that this means that the amount of capital used in production has an important bearing upon the marginal product of labour and, therefore, on wages.

An explanation of what is meant by capitalistic production has already been given. Man may use the original factors of production either directly for the creation of goods of the first order, or he may combine them first for the production of instruments of production which will then be used for the production of consumption goods. In that case, we speak of *indirect*, or *roundabout*, production. The reason for adopting such a roundabout way is a technical one: it is to be found in the fact that indirect production can be more productive than direct production with the same

the only good it needs for want-satisfaction) with the aid of nothing but original factors: land and labour. A suggestion is made that a certain tool should be produced which would, when completed and applied to the productive process, considerably increase the yield. It is found that 20 people would have to devote their labour and natural resources to the production of this tool during a certain period of time, e.g. one year, and they would therefore not be able to produce their own means of subsistence during that time. If we assume that each one of the 100 people normally produces 120 units of the consumption good during the year, it means, clearly, that, with each individual working as hard as before and using the same amount of land, the absence of the 20 who are engaged on the production of the tool would diminish the total output by 2,400 units.

'Feeding' the roundabout process. How, then, can the roundabout way of production be adopted? Three methods are possible. First of all, if a store of consumption goods of 2,400 is present, the 20 people can be fed out of it for the whole time. Alternatively, the remaining 80 may work harder (with diminishing returns and/or the use of more land) and produce an extra 30 units per year each and thereby make up the deficiency. Yet a third way is that, if the remaining 80 reduce their consumption to 96 units per year (while continuing to produce 120), the other 20 could still be maintained at the same level as the remainder of the community. Any mixture of these

methods is possible; and it is also conceivable that, in order to induce the 20 to change over to the production of the tool, they may be offered as large an amount of consumption as they previously got, or even a larger amount. This would involve a still further reduction in the consumption of the others, or a still further increase in their labour.

The essential nature of any of these methods is that there must be available a *supply of subsistence means* which can be used to feed the owners of the original factors (in this case only the workers) while they are not themselves producing consumption goods. Not only is such a subsistence fund necessary, but its size will, given the level of consumption of the community, determine the extent to which production may be made more roundabout. If the subsistence fund is sufficient for a year, then to embark on roundabout production which takes longer than a year would be mistaken: by the end of the year a shortage of subsistence means would make itself felt, and either consumption would have to be reduced or production would have to become direct again. If the error is discovered before the year is up, then a less degree of roundaboutness might yet enable production to be completed. If, on the other hand, a roundabout process taking less than a year is embarked on, the subsistence fund will not be used to the full. But that implies that production is not made to yield as large a return as it might be, since it is made less roundabout than it is possible to make it.

Our example differs in many respects from reality. In the first place, we have assumed that our community begins roundabout production from 'scratch'. In reality, however, the decision is always whether to increase the degree of roundaboutness, whether to begin new forms of indirect production or not, but never whether to initiate roundabout production at all. We have inherited from the past a stock of aids to production, in the shape of tools, instruments, machines, buildings, roads, warehouses, schools, &c., which cause practically the whole of our present production to be of a roundabout character. So what has to be decided from time to time is whether more machines to make machines to make machines are to be made; whether the number of stages of production between the original factors and the goods of the first order are to be increased or diminished; whether the structure of production is to be lengthened or shortened. This fact does not, however, affect our general analysis. The lengthening of the structure of production is still dependent upon the existence of a subsistence fund, and the extent to which it can be lengthened will be determined by the size of that subsistence fund. Nor is any fundamental change necessary when we make our data more realistic by admitting the existence of a variety of goods of the first order which are being concurrently produced and consumed. The process will only become more complex.

But both these modifications make some difference

to the actual use of the subsistence fund. In our simplified example the subsistence fund had to be large enough to support the owners of the factors during the whole time which the roundabout process took. It was only at the end of the year that goods of the first order had matured and were thus able to replenish the subsistence fund which had gradually been exhausted. When, however, many lines of production are running side by side it is not necessary for the subsistence fund which feeds each line to be adequate for the whole period. It is possible for these lines of production to be so synchronized that, in each one, enough goods of the first order are maturing at every moment to ensure the continuance of the roundabout production in the others.

Suppose, for example, that twelve identical roundabout processes, each lasting a year, are taking place side by side, but that each was begun a month earlier than the next one. The result would be that at the end of each month, one roundabout process would be coming to an end and would supply the twelfth part of the means of subsistence necessary to maintain the whole system. But this is sufficient, since the others are at such stages of maturity that each month an additional twelfth part would be forthcoming to take the place of the last one which had been consumed. In reality, therefore, what we first conceived of as a fund, becomes a flow.

Another point emerges from the above. The initiation of roundabout production requires a flow of

means of subsistence; the maintenance of an already existing roundabout process requires that the consumption goods which are turned out should constantly be used to feed the factors engaged in roundabout production. We shall see presently to what extent this necessity is modified by the existence of durable means of production; but we can say already that, in general, unless enough subsistence means are forthcoming for the purpose of maintaining a roundabout process, production will have to be made less roundabout, or return may even have to be made to direct production.

Yet another important difference between our example and reality lies in the manner in which the decision to make production more roundabout is made. We have assumed something in the nature of a collective decision. It is, however, one of our social data that economic activity is individualistic in the sense that there is no central authority, and that the market provides the only link between all the participants. All the decisions involved in capitalistic production, such as the decision to move from the production of consumption goods into the intermediate stages; the decision whether to abstain from consumption and/or to work harder in order to provide the subsistence fund; the decision to maintain a structure of production, to lengthen or to shorten it; all these have to be made by individuals. These decisions will become instances of choice and will be subject to the laws which govern choice. Since they

are, moreover, decisions which can only be effected through exchange, they will be subject to the general laws of the pricing process.

Capital goods, capital and saving. Before we proceed to study the nature of the pricing process in capitalistic production we must introduce a number of terms which are in general use. The goods which are produced by indirect production, other than goods of the first order, are generally called *capital goods*. They can be divided first into *intermediate* products which result from the transformation of the original factors in their passage through the structure of production. They represent, at each successive stage, a greater approximation to the finished consumption good. The second category is generally referred to as *fixed capital*. This is composed of products which are relatively durable, such as machines, and can be used in a series of productive processes.

Distinct from these capital goods, we have capital itself, or *free capital*, as it is often called, which consists of the flow of subsistence means which is destined for the production of capital goods. In other words, capital goods can be produced only if there is free capital. The creation of that free capital itself takes place in the process of *saving*. Saving can be defined as the non-consumption of portions of income. So defined, saving is not enough to create capital; a further activity becomes necessary. This is generally referred to as *investment*. By it the saved income is made available for productive purposes. It is used for the feeding of

roundabout production, or, since in our system exchange is mediated by money, we can say that the purpose of investment is to finance indirect production.

Fixed capital. The existence of durable capital goods raises some special problems. It is true that roundabout production need not involve the production and use of machines. It is conceivable that it would simply result in a larger number of intermediate goods which would ultimately mature. At that point, we should have nothing but goods of the first order and the original factors. In practice, however, roundabout production has nearly always involved the production of machines (in the widest sense) which are still available when the consumption goods have been produced. The production of machines represents in essence nothing more than a particularly roundabout form of production. But their special character lies in the fact that, once they have been produced, production will yield a large return even when comparatively little labour is used.

But when that has been said, the fact still remains that these fixed capital goods are only durable, not imperishable. Sooner or later they will be used up and will either have to be replaced or production will again have to become less roundabout and correspondingly less fruitful. It is true that the necessity of replacement may not become obvious for a long time; it is to this extent that production with the aid of machines differs from roundabout production in general. But supposing machines have been pro-

duced (this, at the time, will have involved a particularly great amount of saving and investment) and they are then used in further production, consumption not only resumes its previous level but even surpasses it since production yields a larger return. A time will come when these machines can no longer be used; they are worn out. Two choices are then open: either all production reverts to pre-machine methods, with a corresponding decline in the level of consumption; or some direct production is undertaken, the level of consumption is even further reduced, there is saving and investment and the machines are reproduced.

This choice can be avoided if, during the lifetime of the fixed capital goods, a fund has been gradually accumulated, which is available at the right time for replacement. Again we see that machine production demands re-creative saving and investment if it is to be maintained in the same way as all roundabout production. Since, however, there may be a certain discontinuity in the production of machines, great variations in the provision of the subsistence fund (i.e. in the rate of saving and investment) are possible. A community may accumulate a large amount of free capital, build up an extensive productive equipment, and then enjoy to the full the increased yield. It may, on the other hand, maintain a sufficient amount of saving to create a replacement fund; or it may even save more in order to be able to add to its capital equipment. Since, to-day, all these decisions are

made by individuals, a great variety of possibilities is present.

We are now in a position to draw a rough picture of the process as it is actually at work. We have different types of production going on side by side, with differing degrees of roundaboutness and maturing at different times. To finance them in the first place, and to ensure their continuance, capital must be invested. This capital is drawn from the incomes of the members of the community and represents their savings. Here again, the entrepreneur steps in. He demands capital and uses what he obtains for the purchase of original factors and of already existing capital goods which he then combines in production. In addition to the markets for factors and the markets for different goods of the first order, we shall, therefore, also have markets for different capital goods, intermediate and fixed, whose prices will be formed according to the laws of exchange.

The price of capital. In addition to these markets, there will also be a market for free capital, on which the supply of savings will meet the demand for capital—the representative of the subsistence fund which alone makes production possible. It is the pricing process in this market which is of special importance.

The price ruling in the market for capital is called *the rate of interest*. How is it determined? We will first of all examine the forces working on the demand side which determine the reserve prices of the entrepreneurs. We know that, under competition, an entre-

preneur will wish to minimize costs, market prices being given. The entrepreneur who wishes to extend production will raise his offers to the factors. This movement must also take place in the case of the initiation of roundabout production, since, in order to obtain the factors, the entrepreneur will have to attract them away from their occupations in direct production by bidding higher prices.

• When production is carried on by means of a combination of factors, the entrepreneur has before him a choice of methods by which to extend his output. If we assume only two factors as variable (keeping the amount of land or raw materials fixed), he can either employ more workers, or make the process more roundabout. Capital, which makes indirect production possible, could be regarded by the entrepreneur as a factor of production similar to land or labour. The pricing of capital must therefore be subject to the same law as that of original factors. That law we know to be the law of diminishing returns or of marginal productivity. The increased use of capital, other factors being fixed, will yield return which will increase at a diminishing rate. Capital and labour are complements. A relative increase of capital will therefore lower its relative marginal productivity or raise the relative marginal productivity of labour. And the price of capital must, as far as the demand side is concerned, tend to equal the value of its marginal product.

The supply of capital. On the supply side, the

provision of capital will be related to its price in the usual way. The supply of capital by the recipients of income is done by means of saving. Their choice will be determined by the place which is occupied on their scale of preference by current and future consumption respectively, and the price which they will ask will express what has been called their 'time-preference'. It is not essential to our present purpose to discover to what extent that time-preference depends on people's income and on their habits of consumption. It can be assumed that in order to draw a larger supply, a higher rate of interest will have to be offered than is necessary to call forth a smaller supply. We must, however, remember that, as in the case of wages, the supply of capital may, in certain ranges, be inversely related to price. Where saving is undertaken in order to provide a certain level of income from interest, a higher rate of interest will enable that level to be reached with a smaller amount of saving. In that case, therefore, a rise in the price will result in a fall, and not in a rise, in the supply. But, in general, time-preference (formed by whatever influences) will determine the reserve prices of the suppliers just as marginal productivity determines those of the demanders of capital. The price will then be formed in the usual way by the competition of the market.

Overhead costs. There remains the question of increasing returns. We have seen that they arise through the existence of indivisible factors. By far the most important examples of indivisibility are to be

found in the case of fixed capital goods. To produce 10 motor-cars may involve the building of a factory and the laying down of plant capable of producing 100 cars. Any increase in the demand for motor-cars above 10 can therefore be met at a decreasing cost per unit. These instances are very frequent to-day, although they become obvious in a somewhat different way, through the existence of what are known as *fixed* or *overhead* charges. An entrepreneur who has laid down a productive equipment of a certain capacity has, thereby, committed himself to a certain fixed outlay, the payment of interest on the capital being the most important. If the demand for the product falls below that of his productive capacity, he finds that the proportion of capital cost which each unit has to bear increases. How can this state of affairs be reconciled with the principle of marginal productivity?

To answer this question we must first of all be quite clear on the position in which our entrepreneur finds himself. He is making losses; his course should be to reduce his production. Why is this not possible in the present instance? Because a reduction in the outlay on the original factors of production, labour and raw materials, cannot help. It is the cost of his fixed capital equipment which is the cause of his loss. But for the fact that this equipment is fixed, that it cannot be retransformed into the divisible units of original factors of which it is composed, our entrepreneur could still reduce his cost. The capital embodied in this productive equipment could be withdrawn from

this enterprise until the point of equilibrium at which diminishing returns begin to operate is reached again: in other words, until the enterprise is working at full capacity. Thus if factors were perfectly divisible, and, in particular, if fixed capital could be retransformed into original factors, increasing returns could not exist for any length of time. At once the tendencies above described would come into operation and diminishing returns would become general.

The existence of excess capital equipment must be regarded as an error which is irrevocable even though it may lead to a changed decision for future production. As original factors are made into capital equipment they become more specialized, and the range of opportunities open to the entrepreneur is narrowed. When he is still in possession of free capital the opportunities are very numerous. Once he has made his choice and has embodied his capital in a specific form, no further opportunities remain. If the choice is erroneous (as shown by the existence of diminishing costs) the entrepreneur will have to carry the loss.

But even with the obstacle presented by the fixed form of capital equipment, the principle of marginal productivity can be found present as a strong tendency. For what will happen in the above example? True, capital cannot be withdrawn. But we have seen that even durable equipment wears out and has to be replaced. Where productive capacity is already too large, i.e. where diminishing costs are in operation, that fixed equipment will not be renewed. If the

enterprise is still capable of raising a replacement fund, that fund will have to be invested elsewhere. In any case, the principle of marginal productivity will constantly tend to be re-established, however slow the process.

II. PROFITS

The constituents of gross profits. Only one aspect remains before the theory can be concluded; that is the remuneration of the entrepreneur. The name generally given to the entrepreneur's remuneration is *Profits*; but, as used in everyday language, profits are a composite form of income. In the first place, a certain part of the entrepreneur's receipts can be regarded as being in the nature of wages. In so far as the entrepreneur is himself exercising supervision over the process of production and is taking charge of organization and administration, he is performing services of a special kind. For that he will be receiving implicit wages. It is difficult to say what proportion of his income will be of that nature; but one thing is certain: the position of entrepreneur is an obvious example of a non-competing group. Entry to this occupation will not be free to all, because it is generally a combination of favourable circumstances such as social position, the ownership of capital, or the knowledge where it can be got, and the ability to get it, which will enable one to assume the functions of an entrepreneur.

When the entrepreneur is himself the owner of the capital used in the enterprise, part of his income must

be regarded as interest. Its share of the total will be ascertained by reference to the prevailing market rate of interest with due allowance for any special features of that enterprise. If the enterprise is a particularly risky one, the rate of interest will be swelled by a special premium for the risk involved. The more the enterprise approximates to one of perfect security, the more will the interest allowance be 'pure' interest.

One of the most important parts of the entrepreneur's profits is that which is in the nature of rent, the producer's surplus which we have already discussed. This explains the existence of marked differences in the rate of profits of individual entrepreneurs. As we know, these differences represent the degree of superiority or inferiority of entrepreneurs in regard to costs. On the margin will be those entrepreneurs who are only just able to carry on. The demand conditions are such that they can obtain a price for their product which only just covers cost and makes it only just worth their while to continue producing rather than to withdraw from business altogether. Above them there are several grades of entrepreneurs, working at different levels of cost, but obtaining some surplus over costs, and finally there are those who are right on top, who are able to obtain a very large surplus income indeed.

These differences in cost may be due to a variety of circumstances. If the entrepreneur has been able to achieve something of a monopoly by means of advertising his branded products, he will have succeeded in

insulating himself from the competition of others; the demand for his goods will be less elastic; and he will be able to charge a higher price or sell a larger amount than he otherwise could. At the same time, the possibility of maintaining steadily a large output may enable him to adopt a greater degree of division of labour, to adopt what are known as mass-production methods, and to cut his costs as well. In this case he will have a more favourable relationship between price and cost than others.

Other factors which make one entrepreneur's income higher than another's are all, in a sense, of a monopolistic nature. The entrepreneur may possess a patented method of production which reduces his technical coefficients. He may possess, in a particularly high degree, the special ability required in his occupation: the qualities of foresight, judgement, initiative. All these would give him some superiority of bargaining, as compared with his rivals, both when he appears as a seller of the product and as a buyer of the factors.

All this explains, on the one hand, some of the elements in that composite term 'profits', and on the other, the marked differences that exist in the level of profits obtained by different entrepreneurs. It does not explain why profits exist at all. It can, of course, be said that profits are simply a positive difference between price and cost which the entrepreneur may or may not obtain. If, for example, the entrepreneur does no part of the administration of the enterprise

himself, but has it done by salaried employees, if he does not obtain any interest, having borrowed the whole of his capital, and he is on the margin and obtains no rent, then all the three elements mentioned above disappear; they are absorbed in cost. Yet the entrepreneur will still obtain an income which will be pure profits. How is that income to be explained?

Theories of profit. The theory of profits is probably that part of economics on which there is least agreement among economists. Only a brief indication can therefore be given here of the kinds of explanation which have been put forward.

The classical economists regarded profits as the reward for the risks which the entrepreneur undertook. Profits were essentially uncertain; the monetary result of an enterprise might be positive or negative. Business enterprise was always risky as a result of this uncertainty. To initiate it involved the shouldering of a considerable risk, but it carried with it also the chance of gain. In the days when the classical economists wrote, business enterprise nearly always went hand-in-hand with the ownership of capital. It was this capital which the entrepreneur risked in his enterprise, and if he was successful, he obtained a reward over and above the interest on his investment.

The difficulty about this theory is that it did not distinguish between the return to capital and the entrepreneur's income. And when the actual conditions of business enterprise altered, it became difficult to apply this explanation. When the growth of joint-

stock enterprise and the development of managerial organization divorced ownership and control, it became difficult to say who bore the risk and who was, therefore, getting profits. Moreover, some risks are of such a nature that they can be adequately allowed for. Actuarial knowledge and the possibility of insurance make it possible to discount many risks.

A different theory of profits is found in the socialist economics of Karl Marx, who built on the economic principles of the classical economists. To him, the only two incomes in the economic system are wages and profits. Labour was capable of producing a *surplus product*, i.e. something in excess of what was necessary to produce the expenditure of effort. But labour was remunerated in accordance with the laws of value, which meant at a price which tended to equal its cost of production. The surplus product was appropriated by the capitalist employers; and it was the value of that surplus product, the *surplus value*, which was the source of profits. These profits were then shared by different classes of capitalists in accordance with their bargaining strength.

Marx also argued that a tendency existed for the rate of profits to become uniform. Competition among capitalists brings it about that the same amount of capital invested produces the same rate of profits. The progress of production means a continual increase in the proportion of fixed capital (or constant capital, as Marx called it) and the amount laid out in wages for the hiring of labour (the variable capital).

Labour is the only source of surplus value; the latter's ratio to the total capital must decline and the rate of profits shows a continual tendency to fall.

More recently, some economists have endeavoured to explain profits by regarding them simply as the remuneration of a special factor of production—enterprise. This may be considered simply as a special form of labour which operates in complex production in the same way as other factors, including capital. It becomes one of the complementary goods which must be used in production side by side with others, and, in so far as the proportion in which it is used can be varied, it is remunerated in accordance with the marginal productivity principle. Differences in profits are explained by invoking the existence of non-competing groups and of rental elements.

At the same time, many modern economists have allowed for the possibility of the exploitation of labour. They have argued that the entrepreneur is in a stronger bargaining position than the worker, since he can withhold his offer, whereas the worker cannot withhold his labour. A big gap exists between the wage-earner and the entrepreneur owing to the fact that modern technical conditions require a very large amount of capital to which the wage-earner has no access. The alternatives of independent production and wage labour, which the classical economists postulated, do not, therefore, exist in practice.

A recent theory developed by the American economist, J. B. Clark, makes profits the result of change.

Under a stationary economic system, with perfect competition, no profits can exist, only wages, interest, and rent. Professor Clark defines a stationary system as one in which population is not changing, tastes remain the same, capital is unaltered, technique does not change, and the forms of industrial organization remain unchanged. In the real world, however, changes in these data are constantly going on; and, in the process, they produce temporary profits to some entrepreneur until a new equilibrium is reached.

Most recently of all, Professor Knight has argued that this last explanation ignores the difference between foreseen and unforeseen change. Some changes in data can be foreseen and their results can be anticipated. Other changes, however, cannot be anticipated; they are due to the uncertainty of the future, in the strictest sense of the word. Professor Knight admits that without change there would be no profits, but it is only because change brings about ignorance of the future that profits do arise. Thus it is not change as such, 'but the divergence of actual conditions from those which have been expected and on the basis of which business arrangements have been made' that causes profits.

The analysis of the way in which such divergence arises is a particularly difficult part of economic study and leads to some of the most advanced problems with which modern economics is concerned. An indication of the quality of the changes themselves is given in the following pages of this book.

PART THREE

THE PROBLEMS OF CHANGE

I. THE QUALITY OF CHANGES IN DATA

Endogenous and exogenous change. The short discussion of profits at the end of Part Two raised the problem of change. In this part we shall examine some of the implications of changes in our fundamental data.

Our analysis of the working of the economic principle was conducted on the basis of definite conditions. We endeavoured to deduce certain general laws of consumption, exchange, and production from premises which we considered to be self-evident. In developing our laws, in particular those which relate to exchange and production, we have been obliged to add, from time to time, certain hypotheses. Thus we have had to postulate the existence of certain scales of values, of certain coefficients of production, of different numbers of participants in the economic system and other data of this kind.

What we did was to make one of the assumed conditions vary while maintaining the others unchanged and observe the reaction of the whole system of relationships to this change. This enabled us to study the *internal* relationship of all the elements in the economic system, supply, demand, price, cost, wages, interest, and to demonstrate their interdependence. It also made it possible to show that there was one

such relationship at which all these elements were in a mutual balance from which there was no incentive to depart; this we described as equilibrium. Our purpose was served, not by assuming any *specific* conditions, but rather by assuming certain conditions to be given at all.

Such mutual dependence would exist whatever the concrete conditions; and it would be possible to conceive of a position of equilibrium within any framework of data. In one sense, therefore, the existence of altered data does not present any special problem. Demand, price, and all the other quantities would still be related to one another in the way indicated in the laws of exchange and production. But the problem of change is to be viewed rather as the adjustment of a given system of relationship to a change in one of the quantities; it does not require the study of the new equilibrium existing after the forces of change have come to rest, but an inquiry into whether such a new equilibrium can be brought about and into the manner in which it is established.

The changes in data with which we are here concerned are those outlined by Professor Clark and mentioned in the last part. They affect population, tastes, technique, the amount of capital available, and the forms of industrial organization. Changes in the population of a community include both total numbers as well as age and sex composition. Changes in tastes are changes in consumer's preferences for goods of the first order. Changes in technique affect the

degree of division of labour and the technical coefficients. Increase (or decrease) in capital implies an increase (or decrease) in saving as against spending, in future demand as against present demand. And, finally, changes in industrial organization affect the number of units at different stages of production and exchange.

In what respects do these changes in the economic quantities differ from those studied in our theory? We have, for example, already contemplated changes in demand, and endeavoured to trace out their effects; in what way would changes in consumer's preferences be different? The kind of changes in demand with which we have hitherto been concerned have only been changes due to alterations in the other economic quantities, but not changes due to an alteration in the scales of preferences of the consumers. We have studied how, given a certain scale of values, price and demand will be related. What we are concerned with here, however, is the effect of what we might call an autonomous change in demand. We ask the question, how will an existing system of economic relationship react if the demand of consumers changes through 'outside' influences?

The same is true of the other changes. A change in technique, for example, will affect the technical coefficients. We ask in what way will supply conditions change; how will the price of the product react? What will be the proportions in which the factors will be used? How will their prices move? We

study, then, the way in which the economic system receives these changes and adapts itself to them. This is the problem of economic development.

Five sets of problems. The approach to it is difficult. There are five different sets of questions we might be asking. First, how do these changes arise? Can we say anything about their causes? Is there any regularity in their appearance? Can we formulate laws of change? Second, we might concern ourselves with the problem of economic equilibrium. Will it be possible? What are the conditions for its realization? Third, we might inquire into the effects of these changes on individuals or groups within the economic system. Fourth, we might ask whether these adjustments ought to be allowed to go on unchecked. Should anything be done to modify the self-regulating character of the system? Should society intervene if it considers that the working of the mechanism tends to produce bad results? Finally, it might be debated whether the changes themselves should be given free play. Clearly, it might be urged that, as an alternative to the interference with the readjustment of the economic system, the causes which make the readjustment necessary (the changes in the data) might themselves be removed.

The philosophy of change. Modern economics can say little about the first set of questions. It may be possible to enunciate general laws of change, but it is beyond the competence of the economist of to-day to do so; the knowledge of the historian, sociologist, and

philosopher is required as well. True enough, the economist can make some contribution towards answering inquiries into the causes of change, for some of the causes lie within the economic system. A state of affairs which makes it necessary to reduce costs may stimulate inventions. A sudden change in demand may, when there is a large fixed capital equipment, give rise to a tendency for monopolistic combination and thus alter the forms of industrial organization.^{*} It may also stimulate the growth of advertising (in an effort to 'attach' demand) and thus alter the tastes of consumers. The growth or decline of capital will obviously be partly dependent upon the size of the income of the community and its distribution. And the size of the population itself may, as we shall see, be influenced by economic conditions.

But with many changes there are other causes to be taken into account. The political system will have an important influence on industrial organization and on the incomes of the community. Saving and spending will be affected by anticipations of the future which include not only economic but other circumstances as well; the fear of war, for example, will act as a deterrent to saving. Tastes may be influenced by fashions for which no *direct* economic reason can be found. Invention, too, may not always be directly traceable to the stimulus of reward. And population will vary as the result of a complex mixture of economic, social, and biological causes.

Opinions will differ as to the weight to be assigned

to these different factors which jointly cause change. Systems have, indeed, been built up round a central explanation. Sometimes it is a psychological one, as in the case of Professor Schumpeter's theory which makes economic development depend upon the leadership of the entrepreneur, his urge for innovation, his impulse to fight. Often an explanation is put forward which is termed an economic one. This is, to-day, perhaps the most widespread of all. Economic developments are regarded as the ultimate causes of all change: political, social, cultural. This kind of interpretation of human development colours much of modern sociological work, and it is met with in mild or extreme form in nearly all the writings of contemporary historians.

Derived largely from Marx, it is generally stated much more crudely than by him. Marx adopted a view which was common (though generally implied) among classical economists. The original cause of change was found in the development of the forces of production, the increase of man's mastery over nature. To exploit them to the full, man organized his society in a certain way. He built first an economic structure; by this Marx understood a certain system of productive relations: relations, in terms of property in factors of production, between different groups of members of the community. Upon that structure is erected a certain political system, a body of laws and customs and even of art and literature, designed to help to maintain the economic structure.

The productive forces continue to grow and, from time to time, they come into conflict with the economic structure, which has then to be changed.

This view of historical development as, indeed, any other comprehensive philosophy of history, draws in factors of explanation which do not seem to lie within the field of the economic theory here presented. Most modern economists have therefore declined the task of explaining change itself, confining their subject within narrower limits: a study of the implication of the relative limitation of resources. If economics is defined in this way, the second set of questions definitely falls within its province. So does the third, for, from a study of the reaction of the system to change, there inevitably follow certain conclusions as to the effects of change on individuals or groups.

The problems of policy. Nowadays, economists are, as a rule, hesitant to claim special competence in the study of the questions under the fourth and fifth headings. These are the questions of policy. Economists claim that to decide what should be done is a question for all citizens according to their own judgments. All that the economist can do is to discover the results of change if no intervention takes place and to show the consequences of intervention. Having done that, he believes that he will have a clear conscience; the citizen will decide 'with his eyes open'.

Opinion on the exact sphere of competence of economists, particularly in relation to the problems of policy, is much divided. Some economists claim

that the knowledge of the consequences of certain policies which the economist alone possesses puts him in a unique position. He cannot hope that others will be as fully conversant with the facts. He is the jury with the evidence before him; the rest are only the readers of the newspaper reports. Other economists, while admitting their special knowledge, would argue that it would be presumptuous on their part to impose their views of what is ultimately desirable on the community as a whole; for any policy of intervention must affect some section of society if not the whole of it, and nobody can be a better judge of a person's interest than that person himself. The economist, they say, is not the jury but the expert witness.

In view of this conflict of opinion, and in view of the complex nature of the problems involved, policy is not treated in this book at any length. In the following pages, however, an indication will be given of the kinds of choices of policies which may arise: their implications can only be studied on a more advanced level. But one further point must be mentioned: it is advisable to treat the pronouncements of economists upon questions of policy as being based, not only upon their special knowledge of the implications of particular policies, but also upon their own predilections; and these deserve no more respect than those of the next man. The economist should command most attention when he makes it perfectly clear what part of his statements represents expert opinion and what part wise counsel.

The effects of change and the problem of adjustment. It is, then, to the second and third sets of questions that we turn. What changes in economic relations will be necessary in order that a change of data may be, so to speak, 'absorbed' into the system? And, secondly, what do these secondary changes involve as far as the relative position of different individuals is concerned?

We cannot, within the compass of this book, deal with all the possible changes of data or even with any one of them at length. But some of the more important must be briefly examined. We shall find that their effects are changes in the relation between supply and demand, either of different classes of goods or of different classes of factors.

2. POPULATION

Malthus. One of the most important of the long-run changes which any developing economic system experiences is an alteration of population. The population problem is one of the earliest to be studied by economists in a systematic manner. The most important of the earlier discussions of the subject was written by Thomas Robert Malthus, and published by him, in 1798, in his famous *Essay on the Principle of Population*. This work, contrary to popular opinion, is less important because of its share in contemporary economic thought than because of the effect it had in starting a prolonged discussion on the subject. In actual fact, Malthus, supplemented by his followers, was con-

cerned both with the nature of changes in population as well as with their effects upon the economic system.

Malthus's own part, at first, at any rate, was almost entirely confined to the former problem. His conclusions are stated in three propositions:

1. Population is necessarily limited by the means of subsistence.
2. Population invariably increases where the means of subsistence increase unless prevented by some very powerful and obvious checks.
3. These checks, which keep population down to the level of the means of subsistence, are moral restraint, vice, and misery.

Basing himself on observations of population increase and of the growth of agricultural production, which he expressed as being in geometrical and arithmetical progression respectively, Malthus concluded that there was always a tendency for population to outrun the available means of livelihood. This could be avoided by positive or preventive checks. The former consisted of famine, wars, excesses, &c., which destroyed the surplus population. The latter consisted of moral restraint, i.e. 'the restraint from marriage which is not followed by irregular gratifications'. (Malthus, it may be noted, did not contemplate the development of contraceptive practices.)

Malthus's followers among the classical economists sought to make this law of population rest on the law of diminishing returns. The difference in the ratios of

increase of food and population, they argued, were due to the operation of this law. An increase of labour on the land was bound to produce a diminishing increase in yield and, unless the checks came into operation, over-population would set in. Once this was accepted, many important consequences followed. Economic progress was seriously menaced by the danger of over-population, and nothing but the exercise of that moral restraint of which Malthus spoke (but in the efficiency of which he appeared to have little confidence) could overcome the danger. The poorer classes of the community, in particular, should be enjoined to exercise prudence; and the later Malthusians drew the conclusion that since an increase in the wealth of the working class would lead to the relaxation of the checks it would defeat its object. Charity, both private and public, was viewed with suspicion, and it was argued that any increase in wages which some workers might obtain would only lead to an increase in population, to a rise in the supply of labour, and to a fall in wages.

It was not difficult for later writers to disprove Malthus's contentions. The facts of economic development during the years subsequent to the *Essay* were alone enough to prove that the danger of over-population was illusory. It was pointed out that agricultural production had increased by leaps and bounds and that the food supply of the world had not only not lagged behind population growth, but had even been sufficient to maintain a very much

larger number of people on a very much higher standard of living.

But even with regard to population itself, recent observations seemed to disprove the tendency for numbers to increase whenever means of subsistence increased. It was pointed out that the relation between means of subsistence and numbers tended to be the reverse of that assumed by Malthus. Among the poor, large families are very much more common than among the wealthy. Every child is regarded as a potential bread-winner and as a support in old age; marriages take place early and families are large. The wealthier classes, on the other hand, will be anxious to maintain a certain social status and standard of comfort; they will wish to provide well for their children; and they will, therefore, marry late and restrict their families. The discovery and spread of knowledge of contraceptive devices and the technical advances in agricultural production seem thus to have falsified Malthus's prophecies. The prospect of stationary or declining populations in many countries, particularly in western Europe, has even led many people to express a fear not of over- but of under-population.

On the other hand, the appearance of surplus population, in the shape of unemployed, has recently led many people to revive the 'Malthusian devil'. They express the fear that, at last, population is beginning to overtake the means of subsistence (at any rate, if these are conceived of in the sense of a conventional

minimum standard of living). The rapid development of productive technique in the nineteenth century, they argue, and the growth of birth control have only postponed the evil day; and diminishing returns are again operating.

Modern economics rejects both these arguments in the form in which they are generally put forward. It is true that Malthus's prophecies have been falsified, but those who point this out are apt to ascribe it to a fault in the law of diminishing returns. It can, however, be shown that the law of diminishing returns was misinterpreted by the followers of Malthus and that it could not, in its modern form, be used as a basis for a prophecy such as Malthus's.

Our discussion of diminishing returns in Part Two of this book gives us the clue to this misunderstanding. In the first place, the contemporaries of Malthus regarded the law as applicable to agriculture only, while they believed that increasing returns were operative in manufacture. It has already been shown that this view is wrong and that the law of diminishing returns is valid for all complex production. In the second place, the law of diminishing returns refers essentially to a condition of stationary equilibrium. It states, with regard to population, that there is, at any time, a certain size of the population which, if exceeded, would yield an increased return at a diminishing ratio. It does not mean that if production is not big enough to use the other existing factors to full capacity, any increase in population will not in-

crease production possibly even more than in proportion.

Nor does it state that an increase in population throughout history would lead to diminishing returns, for in the course of historical development the other factors do not remain unchanged. What the Malthusians ignored and many of their modern critics continue to ignore, is the fact that the law of diminishing returns refers only to a given state of productive technique and was not designed to explain the process of changes in technique. The trouble thus arose through the application to conditions of change of a law which was designed to elucidate the relations which exist in a situation of stationary equilibrium.

✓ The theory of the optimum population. The approach by modern economists is different from that of the Malthusians. It is concerned with the relation between the size of the population and productive efficiency, or, in other words, the amounts of all the other factors which co-operate with labour in production and the technical coefficients of production. From a consideration of this relation arises, as against Malthus's maximum population, the concept of an *optimum population* consistent with a given amount of means of subsistence. The optimum population is that population which, with given amounts of other factors (including land, productive technique, capital, and organization), will produce the maximum product. Any increase or decrease of the population above or below that size will diminish the product.

It is only with reference to this optimum size that economists nowadays speak of over-population and under-population; both imply a total of wealth for the community lower, on account of the size of the population, than it could be made under the existing conditions.

This optimum is not a size fixed for all time: it is determined only by reference to all the other data of the economic system. It is constantly changing with the development of technique and the growth of capital; and it is extremely difficult to say whether at any given moment a particular area is over- or under-populated. Guesses may, of course, be made; but the exact determination of the optimum level is a question not only of the statistical analysis of all data, but also unanimity of opinion on what constitutes the maximum total product. What is often regarded as evidence of over-population is not necessarily the result of a population in excess of the optimum. It may be due to quite different causes, which must be carefully distinguished from the size of the population itself.

Again, it is important to distinguish between the optimum population of the world as a whole and that of specific areas. It is true that it is best to regard the whole world as the economic system to which our economic laws refer. But the world is divided by national frontiers; and different national communities will be concerned with the relation of their own population to the production of wealth in their own area,

rather than with the relation of population to world production. It is possible for such national areas to be over-populated or under-populated in relation to others, if, other things being equal, they produce less wealth than is produced in the other areas.

The effect of population changes on the remuneration of labour is by no means easy to determine. The old post-Malthusian conception that an increase in population, by increasing the supply of labour, tended to reduce wages is no longer adequate. An increase of population may take place without surpassing the optimum, either because population was previously below that optimum size or because, together with the increase of population, capital, technical knowledge, and other factors have altered in such a way as to raise the optimum. In either case the product will increase, possibly even at an increasing ratio. But while the product per head of the working population is going up, the share of labour in relation to that of other factors may decline (it certainly will when other factors remain unchanged and the supply of labour increases up to the optimum). It is impossible to say which movement will be the stronger, the one tending to raise, or the one tending to lower, wages.

3. TASTES

Three types of changes in tastes. As for consumers' preferences, we must carefully distinguish between three kinds of changes. There are, first of all, changes the causes of which cannot be easily

ascertained; these we might call *autonomous*. Then there are others which are themselves the results of changes in population, capital, productivity, &c.; these, for want of a better word, we might call *repercussive* changes in consumers' demands. And, finally, there are changes in tastes which are deliberately brought about: they are *induced*. The results on market relations of all three types of change are similar. As we shall see, they all involve a change in the proportions in which different goods are demanded, with consequent repercussions on the demands for the factors of production. But we shall also find that the implications of these changes are not always the same. Our attitude to their results will vary according to the quality of the original change itself, and so will our views of policy.

Autonomous changes. Autonomous changes are not to be regarded as those which have no cause at all. They are only called autonomous because, from the point of view of the economist, their causation is a matter of some speculation. It is not easy to find examples for such changes. Probably those caused by the caprice of fashion are the nearest approximation. The change-over in women's fashions from long hair to short, from long skirts to short, and vice versa, often possesses the quality of autonomy as we have defined that term. It will affect the demand for dress materials, for hair-nets and hair-pins, and many other goods; and it is not always possible to say that such changes have been deliberately engineered in the

of the population and in its age and sex composition will alter the demand for different goods. A rise in the number of the school population in proportion to the total population will clearly influence the structure of demand in an entirely different direction from that occasioned by a rise in the proportion of the older age-groups who have retired and live on pensions.

Technical improvement and increase in wealth will also affect the proportions of demand. Most of the striking changes in consumers' demand during the last 150 years can be traced to these causes. Improvements in methods of lighting, for example, have diminished the relative share of lamp oil in people's expenditure. Transport developments, caused by growth of population and, in their turn, making it possible to absorb that growth, have greatly increased the proportion of income which has to be devoted to transport. The invention of the internal combustion engine has affected not only the demand for other means of locomotion but the demand for all other goods as well: food, clothing, housing, and entertainment. Improvements in household grates have tended to reduce, relatively, the demand for household coal; and the spread of electric heating will tend to reduce it still further.

It is difficult to know whether, in the sequence of human progress, it is the want which precedes the possibility of meeting it, or vice versa. The wants are certainly always there—in a latent form—and technical development which makes it possible to satisfy

them will make the wants effective. But for this reason the line of demarcation between changes in demand which are the inevitable accompaniment of changes in data, and those which are deliberately induced, is difficult to draw. The possibilities of new satisfaction have to be made known before a want can be expressed. This leads to the growth of channels of communication and to an increase of advertising in all its forms. Once, however, the machinery for making new possibilities known has been developed, it can be used for the purpose of influencing consumers' choices for different ends.

Some of the most important changes in demand are those which are caused by changes in wealth. One of the earliest studies of the connexion between an increase in income and the proportions of expenditure was made in 1857 by the German writer Ernst Engel and embodied in what has since been called *Engel's law*. It consists of two propositions: first, that the poorer a household is, the greater will be the proportion of its income which it spends on food; second, that this increase takes place in geometrical progression. A somewhat similar law was developed by Schwabe, in 1868, in relation to the proportion of total expenditure which is devoted to housing. It states that the wealthier a family is, the greater will be the total sum spent on rent, but the smaller the proportion of total expenditure.

The quantitative measure entailed in Engel's law would be difficult to uphold. It is essentially an

induction and limited by the time and place of the original data. But the general quality of the conclusions of Engel and Schwabe have been confirmed and elaborated by subsequent observations. It has been found that a rising income generally leads to a rise in the proportions spent on clothing and drink, but only up to a certain point. After that they begin to decline once more. The proportion taken up by rent declines up to a point and then rises again. The proportion of food declines steadily; but within the food group important changes take place. Rye bread disappears and wheaten bread takes its place; the proportions of vegetables, fruit, and dairy produce rise; and there is a general increase in the consumption of meat. There is, in general, a movement away from more primitive and inferior articles of consumption to more refined and more varied ones. And every increase in income is accompanied by a more than proportionate rise in the expenditure on entertainment and education.

It is not possible to apply these conclusions without modification to the economic system as a whole. But it does appear probable that any general increase in wealth is accompanied by some such movement as that described above. In Great Britain, for example, there has, in recent years, been a marked increase in the consumption of eggs, milk, poultry, sugar, and chocolate, to mention only a few goods. Part of the changes are, no doubt, due not so much to changes in tastes (whether as the result of changes in income or

for other reasons), but are bound up with changes in the relative prices of goods. But even when allowance is made for the latter factor, there is evidence to suggest that there has been a transference of demands.

2. **Induced changes.** The changes in demand which come under our third heading are on quite a different footing from the preceding ones: they are brought about through the exertion of a deliberate influence on the creation of consumers' preferences in the interests of particular groups of producers. Examples of this type are becoming more and more frequent in modern industrial practice, and they are all based on the use of new methods of salesmanship and advertising. They are directed towards establishing the demand for a particular product and to maintaining it once it has been established.

The use of such practices can be said to arise out of the concurrence of technical improvements which involve large fixed capital equipment and the possibility of competition. Without these practices, the adoption of such technical improvements would always lay the entrepreneur open to the risk of falling demand with the consequence of rising costs. Such a situation, as we have seen, could lead to considerable losses through competitive price-cutting; and it is to avoid these that an attempt is made to influence demand by other means.

Instead of treating demand as a datum, to which he has to adjust his price and output, the entrepreneur will fix output in accordance with the costs dictated

by the technical conditions of production, will fix his price, and will then endeavour to make demand fit his data. His product will be branded; it will generally have a fixed price; and advertising will be resorted to in order to create in the minds of consumers a predisposition to buy that product at that price as against both different products and against similar products of other producers which might be offered at lower prices. The consumers' demand for the particular product is to remain at least stable, regardless of any change in other factors; and, whenever possible, advertising will also be resorted to in order to give the product a larger share of the consumers' total expenditure. In essence, then, demand is influenced in order to abolish competition in the sense in which the word is normally used and in which it has been employed here.

The process described is not a new one. In a minor way it has always existed through the building up of what is known as the *goodwill* of an enterprise. A manufacturer or merchant has always endeavoured to collect round him a circle of buyers who will 'stick' to him, and who will continue to prefer his goods to those of another even though there might be a slight difference in price. But such goodwill could seldom be regarded as permanent; it was always subject to the vagaries of the market. The insulation of the enterprise was strong enough to withstand only minor shocks.

But with modern methods of mass-production and

salesmanship, goodwill is no longer exclusively a question of personal loyalty to be maintained even at a certain cost. It becomes a much more impersonal process of building up preferences and maintaining them. It is a question of forming a habit, unconscious, and, therefore, all the more effective. Rivalry among entrepreneurs continues; but it takes on a somewhat different form. It shows itself not so much in price competition as in the continual, even if slight, variation of the product, and in the search for more effective advertising.

The consequences of such developments are very far-reaching, not least those which affect cultural standards. From an economic point of view, however, it might be argued that such methods may be the only means of avoiding violent fluctuations in production which, with our present economic structure and under existing technical conditions, may lead to serious repercussions on the demand for, and the incomes of, factors. To encourage greater varieties and to immobilize, to some extent, the self-regulating price mechanism would then be the price paid for the avoidance of such fluctuations. It is one of the most difficult tasks of modern economic theory to discover the full implications of such tendencies. It may be said here that it is by no means certain, either that fluctuations can be avoided by the means described, or that the price exacted is not too high.

All changes in consumers' preferences, however occasioned, result in transfers of purchases. They

change the proportions of demand, and, through the consequential changes in relative prices, they affect production. An industry which suffers a decline in the demand for its products will find the prices which it can command falling down to the level of costs already incurred. Its profits will dwindle; and the operation of the law of cost will lead to a restriction in the volume of production. This will mean lower prices for the factors employed in the industry. The interest which can be got on capital, the price which can be obtained for raw materials and intermediate goods, and the wages which can be earned in that industry will decline. Less factors will be employed; the marginal capital, land, and labour will be driven out.

An industry, on the other hand, which has an increasing demand for its products will experience the reverse set of conditions, prices will rise, and so will profits. Production will expand and this will necessitate the drawing in of units of capital, land, and labour which were previously extra-marginal. Wages and interest will rise, and raw materials and intermediate goods will be able to command higher prices in this industry.

Both these movements will result in a tendency for factors to be redistributed. There will be a strong incentive to withdraw capital and labour from industries with a declining demand; and even where this is not possible (as in the case of large fixed capital equipment), the tendency will be for fresh capital or labour

not to enter the industry. An expanding industry will attract factors. The higher rates of remuneration will act as incentives; and capital and labour will tend to move into the industry.

We find, then, that the reaction of the economic system to changes in demand consists of a redistribution of resources among different occupations away from those with a declining demand and into those with a rising demand. If and when such redistribution of resources has taken place, equilibrium will be restored. The real problem arises, however, through the resistances which this redistribution encounters. But before we discuss these, we must first examine the effects of technical changes.

4. TECHNICAL PROGRESS

The effects of inventions and the productive improvements to which they lead have been subjects of continuous study since the beginning of economic science. The problem of the relation between wealth and technical progress occupies not only a long chapter in the history of economic theory, it is also one which has on many occasions during the last 150 years been of great practical importance.

Early reactions. The earliest reactions of both theory and practice produced an attitude antagonistic to the technical improvements. The first days of the introduction of machinery saw a series of violent conflicts. The craftsman regarded every new machine as a rival with whom he could not possibly compete.

The effect of the machine was to cheapen production and lower prices. The craftsman became an extra-marginal producer and there was nothing else for him to do but to cease being independent and to become a wage-earner. This change-over was viewed with great hostility; and the development of the factory system and of machine production was accomplished by a life-and-death struggle against handicraft production.

But even when the independent artisan had been reduced to a position of very little importance in the productive scheme, conflicts did not cease. Now it was the wage-earner who objected to the further introduction of the machine. The machine replaced human labour and led to a decline in the demand for workers which resulted in unemployment. The workers banded themselves together to resist this movement. The Luddites, the machine-wreckers, were the expression of this workers' struggle against the competition of mechanical contrivances.

The fears of the working class were shared by some employers, those who found it difficult to keep pace with technical development; and the theory grew up that mechanical inventions were detrimental and should be prevented. By regulation, by the creation of a hostile public opinion, and by the preservation of superstitions, the development of inventions was made difficult. The early history of modern technique shows many examples of the obstacles which had to be overcome before a new invention became widely adopted.

The forces of resistance could not, however, hold out for long. The output of new inventions in the hundred years 1750-1850 was very rapid, but the cheapening of production went hand in hand with the continual opening up of new markets and there was little difficulty in reabsorbing the labour set free by the introduction of machinery. Individual hardships existed and were, at times, very severe; but the eventual long-run effect of mechanical development was beginning to be regarded as advantageous.

Classical economics first expressed this view in a systematic theory. It treated the introduction of machinery as part of the problem of the increase in the division of labour; and this, as Adam Smith had pointed out, was limited only by the size of the market. Say, the French counterpart of Adam Smith, formulated a *law of the market*, according to which every supply of goods involves a demand for goods and vice versa. An increase in the productivity of labour resulting from the introduction of machinery must, therefore, by increasing the supply, also increase the demand for goods. Thus, an increase in the demand for labour itself would follow; and any displacement of labour could only be of a temporary character. The hardships it entailed would also be temporary only and would be confined to particular groups of people. They were to be regarded as being the necessary price to be paid for progress.

Ricardo qualified this view by arguing that it was possible for the entrepreneur to introduce improve-

ments in technique without increasing the total product and sometimes even decreasing it. This reasoning was attacked by later economists, particularly by Wicksell. He showed that it was conceivable that, for a time, the introduction of a machine might increase the profits of the entrepreneur while yet reducing the gross product. This would lead to a diminished demand for labour and to a fall in wages. But this movement could not go on indefinitely so long as there was competition. Some entrepreneurs would, no doubt, go over to the new form of production. But sooner or later the fall in wages would make a return to the old form of production more profitable and labour would be reabsorbed. Thus the new form would not completely oust the old form if a fall in wages occurred. The two would exist side by side in such proportions as would make the profits of entrepreneurs the same in either; and the total product would ultimately increase.

The modern view of the results of technical improvements falls under two headings. Under the first is studied the influence of invention on the total product; under the second, the effect on the distribution of the total product, i.e. on the prices of the factors of production. Opinion on the first is fairly uniform and can be stated as follows.

Progress and the total product. Under competitive conditions mechanical improvements will be adopted if they reduce the entrepreneur's costs of production and increase his profits. He will be able

to produce more than before with the same amount of factors, as much as before with a less amount, or less than before with a still smaller number of units of factors. If he adopts the last course, then, as was pointed out above, some labour will become redundant, wages will fall, and after a time a return to the old methods of production and a reabsorption of labour will become profitable and the total product will increase.

If he adopts the first course and employs the same amount of factors, the total product will go up. If he adopts the second course, and this is the most likely in the first instance, some labour will be set free and will become available for other purposes. As soon as these redundant units of labour have been reabsorbed, the total product must again increase. Thus in any of these three cases the ultimate effect of any introduction of a technical improvement must be an increase in the total amount of the product.

There is, however, an important proviso in what has been described as the most likely first result, i.e. where labour is set free by the introduction of machinery. Only if reabsorption can take place will the total product increase. Whether, and in what manner, that reabsorption will become a fact will depend upon the elasticities of demand and upon the technical conditions of different industries.

If the demand for a product is elastic, a cheapening of that product through an improvement of productive technique will increase the demand for it more

than in proportion to the fall in price. Production will therefore expand and there will be an increased demand for labour in that industry. Labour can be reabsorbed; but it is not certain whether the whole of that previously displaced labour, or less or more of it will be required. A highly elastic demand and a comparatively small initial displacement of labour may result not only in a complete reabsorption, but even in a bidding away of labour from other occupations. An inelastic demand, on the other hand, may leave the industry permanently incapable of using as much labour as it did prior to the introduction of the machines.

But even where the particular line of production is itself unable to reabsorb its unemployed, reabsorption may yet take place in other ways. If, for example, the demand for the cheapened product is highly elastic, but the technical conditions in the industry are such that it does not require as much labour as before, there will be an increased demand for labour in the industry which is producing the new machines.

And if this avenue, too, is closed it is yet possible to reabsorb labour in other industries. If the demand for the cheapened product is not elastic the consumers' outlay on it will be less than before and they will have more to spend on other things. New wants will then be satisfied and the standard of living will rise, while labour can be reabsorbed in the increased production of other goods. It is, of course, conceivable that the increase in the demands for other goods

will be smaller than the improvement of productive technique has made possible. This would mean that the community desires to have a part, at any rate, of its increased wealth in the form of greater leisure, though this desire is likely to show itself also in an increased demand for some goods, e.g. entertainment, thus still giving the possibility of a reabsorption. But even if no increase in the demand for goods takes place, competition in the labour market would bring about a gradual reabsorption of labour, at the same time reducing the hours of labour all round.

Technical progress and relative shares. A separate problem is that of the effects of invention on the relative prices of the factors of production. We have seen that competitive conditions will make all technical progress result in an increase in the total product; but it does not follow that the prices of the factors of production will be increased in equal proportions. Only if technical improvements raise the marginal products of all factors uniformly will the demand for them go up uniformly too. This is not often probable, for most inventions increase the demand for one factor in relation to that for others. The problems caused by inventions which raise the marginal products of factors in unequal proportions are the most interesting. Since, in practice, the most important inventions are those which result in an increased or decreased application of machinery and, thus, in an increase or decrease in the proportions of capital used, the most important changes are those

caused by inventions which substitute capital for labour and vice versa. The former have been known as *labour-saving*, the latter as *capital-saving*, inventions. Dr. J. R. Hicks defines them, respectively, as inventions which increase the marginal product of labour more than that of capital and vice versa. Neither type of invention need reduce the absolute marginal products of the two factors, though it may do so.

There can be little doubt that the labour-saving invention has been the more frequent in history; and the resistances mentioned earlier were, of course, all called forth by the introduction of labour-saving machinery. Capital-saving inventions, too, may occur. Scientific progress may succeed in replacing complicated machinery, demanding a large fixed capital equipment, by a simpler apparatus which requires little capital investment; and the introduction of the 'shift' system can be regarded as being capital-saving. Examples of inventions which reduce the demand for capital in relation to that for labour have not been frequent in the past, though they may well be much more so in the future.

Changes in the supply of the factors may offset the effects of inventions. It is, for example, possible that an invention may increase the demand for labour, but that a concurrent increase in population would prevent the marginal product of labour from rising, and thus make wages rise more slowly than they otherwise would have done, or keep them stable, or even depress them. Similarly, a rise in the marginal

product of capital, due to an invention which made a greater proportion of capital investment possible, may be offset by a simultaneous increase in capital. Changes in the supply of factors may, however, themselves determine the kind of invention which is adopted. A growth in capital (or a decline in population) will tend to lower the rate of interest (*or* to raise wages), and thus make capital investment more attractive. It is then likely that labour-saving inventions (which change the proportions in which capital and labour are used in favour of the former) will be stimulated. A rise in population may result in a fall in wages (or a decline in capital in a rise of the rate of interest), and thus make it more profitable to substitute labour for capital and stimulate the introduction of capital-saving devices.

Two types of inventions. In order to distinguish between inventions due to different causes Dr. Hicks has proposed the terms *autonomous* and *induced* inventions. The former are those which might be described as spontaneous, i.e. they are not traceable directly to any cause operating in the market. It will be noticed that these inventions correspond to the autonomous changes in consumers' preferences already discussed.

Induced inventions, on the other hand, are all those which are due to changes in the relative prices of different factors, particularly capital and labour. It will be noticed that the word 'induced' is here used in a different sense from that which we gave it in connexion with changes in consumers' demands. In order

to preserve a certain symmetry it seems advisable to refer to inventions of the second type as *repercussive*.

We can say nothing about the quality of autonomous inventions and can, therefore, arrive at no general conclusions about their effects. They may be capital-saving or labour-saving; or they may be of that rare variety, inventions which raise the marginal products of factors in equal proportions. There is no reason to believe that, in the long run, they will be of one rather than of another type.

Repercussive inventions may be of two kinds: those which it would not have paid to introduce but for the change in the relative prices of the factors, and those which, even at previously existing prices, would have been profitable had they been available. The distinction is of importance because it is the second kind of induced inventions which may reduce the absolute marginal product of one of the factors. A relative rise in wages or fall in interest will induce labour-saving inventions. If these are of a kind which would have paid even before the change in the relative prices of the factors took place, the absolute marginal product of labour may decline and wages may fall absolutely.

Taking our discussions of the effects of invention on the total product and on the distribution of the product together, we can state as a general conclusion that, in nearly all cases, inventions will result in a transference of factors, that the existence of competitive conditions will make such transfers possible and will cause an ultimate increase in the total product,

but that inventions are likely to alter the relative shares of the total product going to the factors and their absolute shares.

We concluded that both the major changes already discussed, population and consumers' demands, would make a redistribution of productive resources necessary. This, then, appears to be the major result of change on the economic system. What are the conditions for the achievement of such a redistribution? How does it affect different classes of the community? Should it be allowed to proceed unchecked? These are the final problems which demand our attention.

5. THE PROBLEM OF POLICY

Transference. The establishment of equilibrium in a self-regulating economic system requires that fundamental changes in data should lead to a reorganization of the relationships which make up the system. This reorganization we found to consist of a transference of productive factors, in particular of labour and capital, from channels of employment which are becoming relatively less remunerative, either on account of a change in demand or on account of inventions which alter the technical requirements for the factors, into occupations which are relatively more remunerative. The signals, then, for the need for redistribution will be changes in different prices. The market mechanism will then come into play and redistribution will be achieved.

There is abundant evidence that this has largely been the method of absorbing changes in data ever since the growth of the exchange system. This constant redistribution of resources has taken place between different occupations in the same country, between different areas of the same country, particularly if these areas were highly specialized in their occupations, and, even, between different countries.

In all the important countries of the world technical changes, and the changes in demand following upon an increase in wealth, have tended to reduce the number of people engaged in primary occupations, i.e. agriculture, the extractive industries, and many of the old staple trades. There has, on the other hand, been a marked relative increase in the numbers employed in the more refined newer industries, in service occupations and in the distributive trades. Those regions, particularly in the older industrial countries, which had a large percentage of the declining trades have tended to suffer losses of population through the migration of labour into the areas in which the expanding trades are localized. Finally, there has been a marked rise of certain industries in previously less developed countries of the world, with a consequent accentuation of the rise of the newer industries in older industrial countries.

Few of these changes, particularly those of the most recent years, can be regarded as being only in the nature of adjustments of the system to changing data. Many changes in data have not been allowed to work

themselves out unchecked, and have thus led to other readjustments in the system. War, the rupture of exchange relations, the dislocation caused by different policies, have all a share in the redistribution which has taken place. Although fundamental changes in data have been taking place and have led to a redistribution of factors, these causes present the most urgent problems. They cannot be discussed here.

It is clear that this redistribution must greatly affect the relative position of different classes of the community. Labour engaged in declining trades will find itself faced with a diminished demand. Wages will tend to fall, unemployment will appear. The reabsorption of labour, the increase in wealth, and the possibilities of increased leisure which we have found to be the eventual results of progress, will not be able to make up for the immediate worsening of conditions to some classes which any change must bring about.

Transference may take a very long time. If the changes which make it necessary are on a large scale, powerful barriers may appear which impede the flow of resources in accordance with the dictates of the market. Some of these barriers are always present. Difficulties of language, habit, and the costs of transport will often make a geographical redistribution of labour extremely difficult. The differences in the skill and knowledge required for different occupations may make the reabsorption of labour impossible for many years. The development of dairy farming, the growth of the motor-car industry, the relative increase

in the demand for garage mechanics, cinema attendants, and commercial travellers do not all offer equal opportunities to the South Wales coal-miners whom changes in demand have made redundant in their own industry.

‘Rigging’. The hardships caused by changes, temporary though they may appear from the objective point of view of the economist, are real to those immediately affected. Their first reaction is, therefore, of necessity an attempt to check these changes, to prevent the free working of economic forces which would give effect to them. All those whose established position is threatened by changes in the market situation endeavour to stop the market mechanism as far as they are concerned. They attempt to ‘rig’ the market.

Such attempts generally take the form of monopolistic practices designed to increase the bargaining strength of the parties concerned. Price-rings, trade-associations, trade-union policy, advertising—all these are designed for the same end: to change the equilibrium at which the market quantities would arrive in the absence of such practices. All such attempts must work through the market, that is, they must in some way alter the supply, the demand, or the price of goods and factors. It has often been argued that they can never be very strong if they do not receive aid from outside the market. They may succeed for a time only, since monopoly is always subject to limitations, and while they may retard the full effects of changes, they can never stop them altogether.

State intervention. The position is considerably altered when the question arises whether such aid from outside the market should be forthcoming or not; when, in other words, the problem of State policy is at issue. The State possesses powers of compulsion which are far superior to those which any individual or group of individuals may develop. If an individual wishes to raise the price of a good or service, all he can do is to withhold his supply. If there are very many suppliers, this will not be effective. If he controls a large portion of the supply, he may succeed. It is true that inequalities of wealth and the existence of a class of wage-earners cause a very unequal distribution of such bargaining powers. And this, in itself, raises the problem of the use of the political apparatus, the State, for the purpose of intervening in the bargaining of the market, particularly when changes in the data threaten seriously to affect the relative position of different classes.

The State can legislate, the State can enforce legislation; and this gives it the power to affect the market and, to a certain extent, to put it out of action altogether. The question arises: Should this power be used, and, if so, in what way? From economists, the answer has generally been an opposition to the exercise of the State's power, at any rate when it was proposed to use it in order to maintain the established position of a group of people. A general presumption has grown up against State intervention which would confer privileges and would retard or nullify progress.

Free play of economic forces was advocated as the best method of giving quick effect to changes in the data, particularly of consumers' preferences and of improvements in methods of production. Any other policy, it was said, would impede progress and would interfere with the consumers' freedom to choose. Interference thus involves an unjustifiable judgement by authority of the interest of individuals which they alone can assess.

On the other hand, it has been widely admitted that the State may legitimately interfere when the purpose of its action is to preserve the self-regulating quality of the system. It may, justifiably, prevent the spread of monopolistic practices which would destroy the rule of the market. It may also intervene when wider considerations of health, education, and humanitarianism make intervention necessary. It may set up machinery for the control of the quality of foodstuffs, of drugs and medicines; it may pass factory acts or legislate for mines and railways, laying down conditions for ensuring the safety of those who are engaged in them; and it may regulate entries to professions when considerations of public well-being make it advisable to enforce certain standards of qualification.

It has also been recognized that the State may often have to intervene in order to preserve competition by strengthening the position of particular weak groups. Much labour legislation has been based on this principle. In particular, the British Trade Boards Acts

which apply to industries in which the workers are unorganized are an example of the attempt to put two parties to a bargain on a more or less equal footing.

In recent years, however, it has often been urged that the State should very much extend the sphere of its intervention. Rapid changes, some of a fundamental nature, others themselves the results of policy, have more than ever shown that no existing distribution of resources can be regarded as permanent. Capital and labour have often found themselves faced with a rapidly falling demand for their products. In the case of labour, this has meant transference, often on a very large scale, the difficulties of which we have already noticed. In the case of capital, this has involved losses, since the large amount of fixed capital equipment used in modern industry made withdrawal impossible.

The movement for preserving existing conditions grew very strong, and it naturally endeavoured to bring to its aid the powers of the State. Of late, it has been extremely successful in most countries, even in those in which the traditional presumption against the use of the State machine for the creation of privileges was strongest. The most interesting of recent developments is the extended protection afforded by the State to the interests of capital. In the past, the State was often forced to intervene on behalf of labour. The problem of transference in a modern community often presents itself in terms of millions of men. And even though a study of

contemporary economic development would reveal that much of this problem is not due to causes independent of policy, the existence of democratic government makes it inevitable that the weight of those affected should make itself felt and that the State should have to provide aid.

But the openly avowed aim to preserve the claims of capital in the face of changing circumstances is a comparatively new development in State policy. The arguments in favour of such a policy are generally based on the desire to avoid violent fluctuations in production, which would themselves react on the demand for labour. So the State has stepped in. It has insulated industries in one country from the competition of those in another. It has regulated industries at home, has given them a monopolistic structure, has enforced prices, and has regulated entry to trades. It has also given large direct subsidies to different interests, raised by taxing the community.

These developments have roused the strong opposition of those who believe in allowing changes to go on unhampered and to show their full effects on the system. Their strongest argument has been to show that these measures of interference prevent the system from functioning properly, and may, if they go on for long enough, destroy the system altogether and with it the institutional foundation on which it rests. Moreover, it can be shown that production will fluctuate more (as will the demand for labour) where prices are 'rigged' than where they are not.

Those, on the other hand, who have supported the recent developments of State policy have generally held that the readjustments which recent changes have made necessary have been too big to be left to themselves. Some control, they believe, is inevitable in order to alleviate the hardships of the period of transition.

Some have even gone so far as to admit that the self-regulating market economy may not be able to survive if it is more and more replaced by authoritative control. But this they would not regard as an argument against control, but rather as an argument for the replacement of the system by one of complete central planning. Not all of those who are prepared to go so far are aware of the implications of such policy. They often fail to see that the market economy is, as we have seen, the inevitable corollary of certain institutions, particularly of private property in factors of production and of individual enterprise, and that to scrap the market economy would involve a fundamental change in these institutions. Others who see these implications are prepared to advocate far-reaching institutional changes, as well as the replacement of the market economy by one of central planning: in short, the establishment of socialism.

The older middle view regarded the problem of State control as essentially one to be solved not in general, but for each case on its merits. Statesmanship, it is often argued, is essentially a process of assessing the importance of different interests and of intervening when intervention appears, on balance, to confer

greater benefits than injuries. Those who still hold this view believe that the supremacy of consumers' choice can be pressed too far. Many choices, they would point out, are induced; and, in relation to a great many articles of consumption, the consumer is essentially in ignorance and at the mercy of the producer. As regards the individualistic nature of consumers' preference and the absence of any objective standard by which to measure personal satisfactions, they maintain that not only is monetary reckoning all-pervasive, but that there are also widely held uniform views on social justice which do not make it necessary, in practice, to assume a wide disparity in people's wants. Comparisons on the basis of money, while not scientifically precise, can, therefore, be made, and are, in fact, constantly made.

These advocates of occasional intervention should, however, understand that the real difficulty is a political one, in the narrow sense of the word. How is it possible to ensure that the implications of State policy will be fairly presented to the community and will be clearly grasped by it? How can the exercise of undue influence over the political apparatus by particular interests be prevented? Can the State be placed above the contending parties? Can it be an arbiter? Or must it, of its very nature, be the instrument of particular interests? These, however, are all questions of politics, and we must end at the point at which economics abuts on the related fields of social and political study.

6. THE TRADE CYCLE

Far more urgent, in practice, than the long-run changes analysed above are certain periodic convulsions which affect the whole of the economic system. These disturbances are of a different quality from the developmental changes in data which have been examined; and the problems which they create, while they also arise through difficulties of readjustment, are nevertheless much less clear-cut. *Crises*, *Industrial Fluctuations*, or the *Trade Cycle*, as these disturbances have been called, fall really outside the scope of an elementary work. The literature on the subject is very large; and while a certain measure of broad agreement has been achieved, differences of opinion between economists, both on the diagnosis and the therapy of the trade cycle, are still so considerable that an advanced knowledge of economic theory is required for any detailed discussion. The few points which are given below should, therefore, be regarded as indications only of possible lines of inquiry.

The symptoms of the trade cycle. The facts of the trade cycle are comparatively simple. Fluctuations in the level of economic activity have always existed. Catastrophes such as droughts and plagues, wars and invasions, and the sudden occurrence of some of the changes cited earlier, have been responsible for disturbances of the economic machine, with the consequent violent alternations of busy and slack

periods, long before the development of the modern economic system. But the periodic, rhythmical, up-and-down movement of business, to which we give the name of the trade cycle, and which is not in any obvious way traceable to any of these 'natural' causes, is a product of the modern capitalist system; it is a phenomenon of the last 150 years or so.

The symptoms of these alternations of booms and slumps are, briefly, as follows. At certain times, business activity is clearly following an upward movement, although the periods when it is maintained on a comparatively high level are short. Production increases; the factors of production are well employed, even less efficient units of capital and labour having been drawn into the process of production; the remuneration of the factors tends to be high; above all, the profits of the entrepreneur rise.

A sharp rise to which we give the name of boom supervenes and, after a short period of feverish speculative activity, prosperity generally ends in a critical convulsion during which all the high structures previously built collapse. After a time, business settles down to a more or less prolonged period of low activity: the depression. Production figures are low; there is unemployment both of labour and capital; the remuneration of the factors of production is low; profits have shrunk to a very low figure, and many entrepreneurs are carrying on at a loss.

The character of the trade cycle. The two chief characteristics of these movements are that they are

cyclical and that they tend to be *general*. The first implies a distinct rhythm consisting of different phases which follow each other in a well-marked order. Roughly these phases can be classified as (a) upward movement or recovery, (b) boom, (c) crisis, (d) depression. A study of the fluctuations which have occurred in recent history shows that these phases are not always of equal duration or intensity. There are certain secondary characteristics peculiar to each cycle which lead sometimes to an overlapping of phases, sometimes to the possibility of distinguishing additional ones. Nor has it been possible to establish anything like a uniform duration for all cycles. Not only is there a big margin of variation in the length of different cycles, but the statistical data on which estimates of such length are based have been variously interpreted by different economists. And while a duration of between five and ten years appears to mark the cycles of the last 100 years, it is best to insist on the general cyclical quality of fluctuations, rather than to claim any definite time-range for them.

The *general* nature of the trade cycle is its other necessary characteristic. This is a quality which has become progressively more marked as the modern economic system has developed. Nothing but a disturbance of the whole economic machine could create a special problem of analysis and policy, since partial dislocations must, of necessity, be due to special, comparatively easily found, causes. Modern crises

are, therefore, the product of the capitalist market economy of which the development was sketched in a previous section. Not until all trades, all commodities, and all members of the community had been drawn into the system of capitalist production and exchange, and had been linked by the ties of the social economy, was it possible for such general movements to occur. The trade cycle, then, must affect all trades and all districts of a community, though not necessarily with equal intensity.

It is true that in the early days of modern economy this rule was not always fulfilled. But it could generally be shown that when a particular area or a certain branch of production (such as agriculture) remained unaffected, this was due precisely to the absence or weakness of any link between it and the economic system as a whole. To-day, it is very rarely that such exceptions can be found. Not only are all branches of the economic system affected, but no major disturbance occurs without spreading *internationally*. This development is the natural outcome of the progress of international trade, of the establishment of world markets, of the adoption, in recent history, of a single monetary standard by all the most advanced countries, of the international mobility of capital: in short, of the making of the world into a single economic unit.

The all-embracing quality of modern cycles has made their study a matter of prime importance. This study has, in the first place, been directed to a

clear description of the course of fluctuations in industrial activity and has thus been essentially statistical. A mass of data on all the relevant indices such as production, prices of commodities, interest rates, wages, unemployment, and volume of home and foreign trade is being accumulated. It was hoped by some economists that some particularly striking correlation of data might emerge which would suggest fruitful lines for theoretical investigation. What is certain is that it is now possible to present a fairly accurate picture of the main succession of events of the trade cycle.

This picture relies on the transformation which certain market phenomena undergo between the upward and the downward swing of the cycle, and describes its internal development, ignoring, in the first place, the question of the original cause of the movement. Beginning at a point of time when the depression has been of some duration, we find a certain temporary balance between all the parts of the economic machine. Production has settled down to a low level and prices are low; less efficient entrepreneurs have been weeded out; there is a considerable reserve of unemployed labour, and wages are low; savings are held in as liquid a form as possible: there is no incentive to capital investment since the profitableness of enterprise is small; interest rates show, as a consequence, variations as between different markets: being lowest in the shortest and safest, and highest in the longest and riskiest channels.

This general stagnation could, clearly, be brought to an end by some outside event, such as the stimulus to some demands caused by government expenditure (rearmament, for example). But whatever the first incentive, as soon as it appears, the very conditions of stagnation will turn into factors which encourage a recovery. Low prices and low wages will act as inducements to the expansion of enterprise; above all, the abundance of funds potentially available for investments will, at the first sign of a return of enterprise, help in the upward movement. Gradually, production increases. It is a well-established fact that this increase first affects the constructional trades (those producing capital goods), since it is in those trades that the low rates of interest on capital are particularly strong incentives to expansion. Prices rise, though again not uniformly. Raw materials, particularly those used in the production goods industries, are the first to be affected, and others follow. And, finally, a later and non-uniform rise in wages can be observed.

Where the turning-point is reached is still a matter of debate. The most obvious symptoms which indicate that the boom has become speculative and has overshot the mark is, generally, to be found in the monetary sphere. There is a sudden rise in the rate of interest, which at once administers a check to enterprise. Loans are called in, an attempt is made to liquidate investments; there occurs the reverse movement of the one observed before: a general shifting of funds from less to more liquid forms. This leads

entrepreneurs to an equal scramble for liquidity: some projects, particularly those which were planned to mature a long time ahead, will be abandoned, others curtailed. Stocks will be cleared, even at reduced prices. Production will be restricted and workers dismissed, even though the consequent decline in purchasing power will accentuate the fall in price.

This critical movement slowly comes to an end; and when a considerable scaling-down of prices, wages, &c., has been achieved, a temporary equilibrium is established at the new low level. The economic system then continues, for some time, in a state of depression which is particularly marked in the capital goods industries. We have again reached our starting-point with the possibility of a repetition of the whole process.

† Theories of the trade cycle. The foregoing description, sketchy though it is, will have shown in what respects the trade cycle differs from the problems of change with which we have hitherto been concerned, why a special theory of the trade cycle is necessary, and, finally, what factors such a theory will have to take into account. To be fully satisfactory the theory will have to explain the two fundamental characteristics of the trade cycle, namely, its generality (including its increasing generality) and its recurrence. Since general economic convulsions have existed in the pre-capitalist era, when they were the results of natural or political catastrophes, it was natural that

some of the earliest theories of the trade cycle should have relied for their explanations on similar natural events. Wars or failure of harvests were regarded as the initial causes, and it was possible to show that the cumulative movement which followed was made inevitable by the complex interrelations of modern economy. An attempt was made to explain the recurrence of cycles either by some theory which claimed to demonstrate the periodicity of good and bad harvests (such as the famous 'sun-spots' theory of Jevons), or by the denial of any regularity in the alternation of boom and slump. If we deny regularity, it follows that any catastrophic event may set up a cycle which can be perpetuated by the fortuitous occurrence of one out of a large number of possible other impulses.

Explanations of this kind were, however, found to be unsatisfactory. Even though no uniform duration of cycles can be discovered, their recurrence is sufficiently striking to demand special explanation. Good or bad harvests may undoubtedly play an important part in the exact sequence of the phases of a cycle. But it can be shown that the closer linking-up of the world, which was bound to diminish the importance of the size of crops in any particular area, has, at the same time, made crises more general.

The first condition for a satisfactory theory is the recognition of the fact that the trade cycle is a phenomenon of our particular social and economic system. No explanation, therefore, which adopts a 'catastrophic' view appropriate to earlier systems can claim

our acceptance. A theory which forms a bridge between these naïve explanations and the more elaborate modern ones is of a psychological character. It takes into account the fact of private enterprise on which our system rests, and explains cycles in terms of waves of optimism and pessimism which, from small beginnings, gather momentum and lead to a self-engendered rhythm of collective error—first upwards and then downwards. Nearly all modern theories would assign some importance to the spread of particular anticipations throughout the system which is made so much easier to-day by the interconnexion of markets. But the above argument can hardly be regarded as a full explanation either of the initial cause or of the particular relative changes that are present in cycles.

Another theory which goes farther into the real nature of the trade cycle is that which starts from the striking and well-observed fact that the proportions in which production and consumption goods are produced, and their relative prices, undergo violent changes in the course of the cycle. Nearly all existing theories make use of this fact in their explanations. Some economists have argued that the cycle is due to a sudden burst of capital development, a sudden jump in the accumulation of capital, due to the opening-up of a new country, to the discovery of new resources or processes, or to some similar cause. They point to the fact that most cycles of the nineteenth century were associated with some such well-defined expansion in

capital development: railways, the opening-up of South America, the introduction of electricity, oil, &c. A large quantity of capital goods is required in the initial phases of such development; but once this has been produced, demand falls off. In other words, the demand for capital goods is discontinuous; and this discontinuity is the cause of the cyclical movement of business activity as a whole.

A large number of other theories include, or rely exclusively, on an analysis of this disproportion between capital goods and consumption goods. Some of these—generally referred to as under-consumption theories—claim that there is an inherent tendency for capital accumulation to proceed, at times, at too rapid a pace. Since the ultimate purpose of capital expansion must always be to raise consumption, i.e. to increase the output of consumption goods, it is easy to see that there may be certain rates of accumulation at which a discrepancy arises: the market for the products of the increased capital equipment is not able to absorb the supply, and a general decline follows. The exact manner in which these disproportions arise and the quality of the causes which inevitably make capital development outrun the possibilities of consumption are matters of great controversy. The explanations range from simple fallacies which economics has little difficulty in exposing, to very elaborate theories which introduce a variety of fresh factors.

Chief among the additional factors which modern theories take into account is the working of the mone-

banking system. This has created—in a manner which it is impossible to analyse here—money substitutes, credit; and through alterations in the volume of credit and the rate of interest from loans it can exert a powerful influence upon the state of markets.

Since money is also a means of generalizing purchasing power, it becomes a powerful instrument of expressing, and making general, individuals' anticipations. The existence of money and the elaboration of money and capital markets with channels of investment of varying lengths and risks, presents the individual with the possibility of postponing the consumption of income (equivalent to a postponement of the second half of the act of exchange: $G-M-G$) and of keeping the non-consumed portions of income in forms of varying liquidity. We see thus a glimpse, at any rate, of the connexion between the money system and the problems of anticipations, on the one hand, and between it and the general waves of optimism and pessimism which undoubtedly form part of the trade cycle, on the other. In the working of the monetary system may also be found the explanation of the technique by which the sharp turning from boom to slump and the gradual change from depression to recovery are achieved.

Enough has been said to show in what directions a complete explanation must be looked for. It will have to account for the changes in relative prices, interest rates and wages; for the disproportions in the sphere of production; for the conflict between consumption

and accumulation; for the changes in anticipations; and for the role of the monetary system. Finally, and most important of all, it will have to explain the relation between these separate phenomena and the fundamental institutional data of our economic system.

The problem of policy. Very little can be said about policy in relation to the trade cycle; although it must be recognized that there is no more important practical problem facing us to-day than the cure and prevention of crises. Obviously the remedy which is suggested depends very much on the diagnosis which is made. According to the emphasis placed upon the factors enumerated above, policy will rely on measures to stimulate consumption or accumulation; on attempts to check or to encourage the expansion of the capital goods industries; on monetary measures designed sometimes to increase and sometimes to diminish the volume of effective money.

Disagreements among economists concerning policy are far greater than those of analysis. Differences of emphasis in diagnosis, especially since there has been an increase in the eclecticism of theories of the trade cycle, have not been very striking. But when it is a question of suggesting policies, these differences of emphasis become all-important. Some economists maintain that the explanation of the crisis is to be found in the preceding boom. The only way, therefore, to even out fluctuations is to prevent the recovery movement from developing into a speculative boom.

Once, however, it has been allowed to go so far, nothing can prevent the collapse and depression. At that stage, the best policy is to allow the liquidation of the unsound positions created in the boom to be completed as quickly as possible. The introduction of palliatives would only prolong the agony.

Other economists lay more stress on the self-accelerating character of the downward movement once a crisis has occurred. They would plead for measures designed to stimulate demand through the expansion of public works and the maintenance of the purchasing-power of the mass of consumers. Monetary measures are of importance in both these policies: it is claimed in the one that an early restriction of credit will prevent the boom; in the other, that an early expansion of credit will prevent the crisis from turning into a lengthy depression.

It must be clear from what has been said earlier about economic policy in general that the same problems concerning the functions of the State are involved. Monetary policy, it is true, has always occupied an exceptional position. Even the staunchest advocates of *laissez-faire* among the classical economists admitted that State intervention to regulate the issue of money and credit was legitimate. But there are few economists to-day who have faith in the ability of monetary policy alone to remove fluctuations; nearly all the schemes that are put forward involve some additional measure of State intervention. And those economists who believe that fluctuations are

inseparable from the existing social and economic system must be prepared to argue on the much wider terrain of political action. Once again, therefore, we conclude at a point at which economics, as to-day understood, must give way to politics.

CONCLUSION

IT seems advisable to add a few concluding words, not so much as a summary of what has been discussed in the preceding pages, but in order to indicate some of the more advanced problems to which it is necessary to proceed. These problems can be divided into three groups.

There are first the problems of more detailed economic analysis: the development of the elements of pure economic theory presented in Part II of this book. These include refinements of the analysis of equilibrium and a study of alternative theories, as well as more searching discussions of the disturbances to which the system is subject and of the methods of readjustment to change. These problems are very numerous indeed; but a few may be mentioned here.

In the theory of value, for example, there is a conflict between those economists who regard a theory of choice as essential and those who wish to confine themselves to a theory of exchange. The latter argue that it is unnecessary to make complicated assumptions about the nature of people's conduct, assumptions which are not capable of verification. All we need do is to observe people's behaviour in the market and to formulate a theory of price in accordance with our observations. The other school maintains that it is impossible to develop a satisfactory theory of price without the instrument of a theory of choice which

refers to people's conduct even when exchange (an institutional datum) is absent.

In the theory of the market itself, much work has recently been done on exchange and production in the presence of different degrees of monopoly and competition. The theory of duopoly and oligopoly has been much elaborated; and the very complex conditions of the existence of equilibrium in the market as a whole, as well as for each participant, are now being studied. The resistances to changing conditions have also received a great deal of attention, and, in particular, the changes in the degree of competition consequent upon changes in data.

Perhaps the most interesting work is that which relates to the uncertainty of the future, to the existence of money, and to the trade cycle. Here the problem has, roughly, been that of studying the quality of anticipations of the future and their effects on market equilibrium. Changes in anticipations will make themselves felt particularly in changes in the proportions of saving and spending, and, thus, in the structure of demands for different goods and for factors. The possibility has been analysed of general changes in the desire to hold money for future uses and in the effect which these have on the law of the market, which states that a supply of goods always entails a demand for goods, and vice versa.

In addition to these theoretical problems, investigations into the facts of the economic structure have much increased in recent years. A large volume of

information is available on the occupational distribution of the population, on the size of the product, on the productivity of labour, on consumers' outlays, on relative prices, and on changes in all these data.

A study of such factual information is an absolute necessity for the economist. It cannot replace the continual development of the implications of our definition of economics which is the proper field of pure theory. But it will furnish new problems to which the deductive process may be applied; it will suggest new hypotheses; and it will enable theory to be subjected to continual verification. It will, moreover, supply that general background of knowledge of affairs which will help to make the forecasts of the economist more reliable.

Finally, there are the problems of policy. The classical economists made their economic reasoning part and parcel of a general system of social ideas. They were thus able to pronounce without hesitation on practical issues and to take a prominent part, as economists, in the fashioning of policy. Modern economists, as has already been pointed out, have tended to restrict the sphere of competence of economics by making the subject into an instrument for revealing a system of interrelationships which exists under certain assumed conditions and, thereby, to discover the implications of any change in these conditions which may be proposed. The role of the economist is more and more conceived of as that of the expert who can say what consequences are likely

to follow certain actions, but who cannot judge, as an economist, of the desirability of these actions.

When interpreted in this way, economics can contribute little to the ultimate solution of problems of policy. Provided that the issues have been made clear, there is nothing in economic theory which can decide whether any piece of State intervention is desirable or not. And if the implications are understood, the economist cannot even say whether the self-regulating system which he has been studying is superior to any other. This change in the approach to economics is not sufficiently realized by public opinion. The authority of economics is still quite often invoked to support a course of action which is based on considerations about which modern economics has nothing to say. It is important, therefore, that the extent of the competence of our subject should be clearly understood.

Whether the recent restriction of the competence of the economist is itself desirable or not, is perhaps the most difficult and certainly the most fundamental problem of all. There can be no doubt that the process of making economics more scientific has yielded many valuable results. It is also true that further developments along the same lines may be necessary if the economist is to play his part as the expert who presents the community with a clear statement of the implications of different policies.

But satisfying though this more rigorous formal study may be, it carries with it serious dangers. It

may so fascinate the student that he becomes blind to wider issues, or too cowardly to face them. Economics started as a practical subject concerned with the betterment of the conditions of humanity. However austere a scientific discipline it may become, it will always be necessary to put it in its proper setting of social and political ideas.

This necessity has so much impressed some economists that a reaction against the growth of formalism, against excessive neutrality, has set in. This reaction has taken the form either of a revival of Marxism, or, and this is much more noticeable in academic circles, of a combination of rigorous analysis with what may be called teleological assumptions, i.e. preoccupations with particular ends or purposes. At the present time, therefore, we can distinguish, roughly, three trends in economic thought: the Marxist, the formalist, and the teleological.

In a book which is an exposition of 'orthodox' economics little can be said about Marxism: some indication has already been given of the very comprehensive nature of Marxist thought. It has been shown that Marxism is based, in the first place, upon an analysis of the contemporary social structure from the point of view of property relations. From an investigation of the evolution of that social structure Marxism claims to derive certain general principles of the development of society which, briefly summarized, are: the existence of opposing classes (i.e. classes with antagonistic interests due to the difference in their

relations to property in the means of production); the perpetual conflict between these classes and between the class character of society and the developing productive forces; and the periodic social revolutions by which these conflicts are solved. These principles are applied to an analysis of present-day capitalism and yield a number of theories relating to value, exploitation, capital accumulation, and crises which are, however, too complex to be dealt with summarily here.

The formalist school, which has greatly developed during the last sixty years, is split into those who, as has been mentioned, wish to preserve the concept of utility, and those whose desire for scientific rigour has led them to concentrate on the mathematical treatment of measurable quantities and increments. There can be no doubt that it is this latter method which has produced the more imposing results in recent years. The unity of the two sections persists in their common desire to avoid reference to social institutions and to establish laws similar to those of the physical sciences. The unifying principle of their analysis is that of choice, or substitution, which can be regarded as a natural datum; while the unifying principle of the Marxist system is the class nature of society, a social datum.

The third of the present tendencies includes the considerable use of the modern analytical technique, but makes use also of teleological hypotheses, such as the avoidance of fluctuations, full employment, or more equal distribution of wealth. These are put on

a level with other hypotheses constantly used in economics; their implications are worked out; and a much surer guide to policy is then claimed to be available. How far such a procedure is compatible with the continuance of the claim that economics is impartial, it is difficult to say. It is, at any rate, clear that a consistent adherence to this method must again raise problems of politics. In the last resort, therefore, the new teleological school must be judged on its own ground: in accordance with the clarity with which it recognizes the implications, in terms of politics, of the economic policy which it advocates.

BIBLIOGRAPHY

THE following is a short list of books recommended for further reading. Many of them are on a more advanced level than the present book; but it is difficult to find good works on special topics which are not more advanced than a general introduction.

THE SCOPE OF ECONOMICS

The standard work in English is: KEYNES, *The Scope and Method of Political Economy* (Macmillan); but it is a little old-fashioned.

The chapter on Political Economy in MILL's *System of Logic* (Longmans), though also old-fashioned, is still worth reading.

ROBBINS, *The Nature and Significance of Economic Science* (Macmillan), is not easy, but it gives the best statement of the positive approach.

DOBB, *An Introduction to Economics* (Gollancz), shows well the implications of that approach by means of a survey of the development of economic theory.

For those who read French: BOUSQUET, *Institutes de Science Économique*, vol. i (Giard) may be recommended.

CASSEL, *On Quantitative Thinking in Economics* (Oxford), is a statement of the grounds for rejecting the theory of value.

GENERAL WORKS

Of the larger text-books, TAUSSIG, *Principles of Economics*, 2 vols. (Macmillan), though a little out of date, is still the best.

For those who read German, OSWALT, *Vorträge über Wirtschaftliche Grundbegriffe* (Gustav Fischer), can be recommended as an excellent introduction.

Other general text-books are mentioned below in connexion with particular topics.

THE THEORY OF CONSUMPTION

The best statement is to be found in WICKSTEED, *The Commonsense of Political Economy*, vol. i (Routledge).

Part III of KNIGHT, *Risk, Uncertainty and Profit* (London School of Economics Reprints), is difficult to read, but contains an excellent discussion of the problems involved.

More advanced, and largely mathematical, is HICKS and ALLEN, 'A Reconsideration of the Theory of Value', in *Economica*, February and May 1934. The first part, at any rate, should be studied.

THE THEORY OF EXCHANGE

A large number of books exist on this part of the subject. The following is only a small selection of larger works of which the relevant portions should be studied. WICKSELL, *Lectures on Political Economy*, vol. i (Routledge); MARSHALL, *Principles of Economics* (Macmillan); the books by KNIGHT and TAUSSIG already mentioned, and CASSEL, *The Theory of Social Economy* (Benn). For a special part of the subject, WHALE, *International Trade* (Home University Library), and TAUSSIG, *International Trade* (Macmillan), may be mentioned. Recent work on the theory of monopoly is rather advanced, but an excellent discussion of the main problems of duopoly will be found in PIGOU, *The Economics of Stationary States* (Macmillan).

THE THEORY OF PRODUCTION

WICKSELL's *Lectures* and KNIGHT, *Risk, Uncertainty and Profit*, provide the best comprehensive treatment in English. CASSEL, *The Nature and Necessity of Interest*, deals with capital, and there is a very good German work on Capitalistic Production: STRIGL: *Kapital und Produktion*

(Springer). A useful short introduction to wages is ROBBINS, *Wages* (Jarrolds). The most complete and up-to-date (though advanced) statement of the Marginal Productivity Theory is to be found in HICKS, *The Theory of Wages* (Macmillan). DOUGLAS, *The Theory of Wages* (Macmillan) combines very well theoretical and statistical studies. An admirable critical discussion of the problem is DOBB, *Wages* (Nisbet). TAUSSIG, *Wages and Capital* (London School of Economics Reprints), is an important work on the connexion between the theory of wages and the theory of capital. ROBINSON, *The Economics of Imperfect Competition* (Macmillan) and CHAMBERLIN, *The Theory of Monopolistic Competition* (Milford), are advanced, but very important. An excellent introduction to the problems of production under conditions of imperfect competition is KALDOR, 'The Equilibrium of the Firm', in *Economic Journal* (March 1934). And indispensable for a more advanced knowledge of the theory of production is SRAFFA, 'The Laws of Returns under Competitive Conditions', in *Economic Journal* (Dec. 1926).

THE PROBLEMS OF CHANGE

Most of the larger works deal with these. The following might also be mentioned: WRIGHT, *Population* (Nisbet); ROBBINS, 'The Optimum Theory of Population', in *London Essays in Economics* (Routledge); GREGORY, 'Rationalisation and Technological Unemployment', in *Economic Journal*, December 1930; ROBBINS, 'The Economic Effects of Variations in Hours of Labour', in *Economic Journal*, March 1929; FISHER, *The Clash of Progress and Security* (Macmillan).

There is a very large number of descriptive works and of books on monetary problems, on the trade cycle, and on Marxism. Here are a few: ASHLEY, *The Economic Organisation of England* (Longmans); CROWTHER, *Ways*

and Means (Macmillan); ALLEN, *British Industries and their Organisation* (Longmans); ROLL, *About Money* (Faber); ROBERTSON, *Money* (Nisbet); LEAF, *Banking* (Home University Library); MACFIE, *Theories of the Trade Cycle* (Macmillan); STRACHEY, *The Nature of Capitalist Crisis* (Gollancz).

INDEX

- Acquisition, 26.
- Advertising, 159, 164, 196-7, 206, 223-5.
- Agriculture, 150-3, 211-14.
- Allen, R. G. D., 66, 107.
- Anticipations, 100, 201, 206, 255, 263.
- Average product, 151.
- Bilimovič, A., 34.
- Booms, 248, 254, 259-60.
- Capital, 187, 200, 203-4, 216, 217, 226.
 - and the State, 244-6.
 - demand for, 190-1.
 - fixed, 187-90, 199, 226.
 - goods, 187, 256-9.
 - market, 258.
 - price of, 190-2.
 - supply of, 191-2.
- Capitalistic production, 20, 179-95, 250.
- Chamberlin, E., 86.
- Change, 200-1. Part Three.
 - autonomous, 204, 218-19.
 - induced, 217, 223-5.
 - philosophy of, 205-8.
 - repercussive, 218, 219-23.
- Choice, 13-17, 43-5, 116-17, 187, 246, 262.
- Clark, J. B., 200, 203.
- Combination, 158.
- Communal scale, 91-6.
- Competition, 26-8, 104-6, 164, 199, 223, 266.
 - cut-throat, 157-60.
 - bilateral, 76-81.
 - imperfect, 84-7, 162-6, 197, 223-5.
 - perfect, 83-4, 157, 161-2, 165-6.
 - unilateral, 69-71.
- Complements, 96, 106-10, 200.
- Consumer's surplus, 79-81.
- Consumption, 34-8, 192, 217-27.
 - goods, 48, 49, 117-18, 255-6.
 - in the isolated economy, 59-62.
 - theory of, Part Two, Chapter One.
- Cost, 128.
 - and output, 153.
 - — — in perfect competition, 161-2.
 - — — in imperfect competition, 162-4.
 - average, 161-2.
 - comparative, 139-43.
 - constant, 150.
 - of labour, 172-3.
 - decreasing, 152-7, 194.
 - increasing, 150, 154.
 - joint, 144-9.
 - law of, 128-30, 136-7, 145.
 - marginal, 137-9, 142, 145, 146, 161-3.
 - opportunity, 133-5, 194.
 - overhead, 192-5.
 - price and, 128-33.
 - — — in perfect competition, 161-2.
 - — — in imperfect competition, 162-4.
 - special, 147.
- Cournot, A., 85-6.
- Credit system, 257-8.
- Crises, 247-61.
- Demand, 71-2, 94, 202-4, 225-7, 257.
 - derived, 117-18.
 - elasticity of, 102-6, 148-9, 158-60, 197, 223, 231-2.
 - for capital, 190-1.
 - for labour, 172-9, 234.
 - for money, 98-100, 263.
 - schedule, 101-2.
- Desirability, 53.

- Diminishing returns, 150-5, 160-1, 181, 191, 211-15.
 Diminishing utility, 54-7, 154.
 Discrimination, 73-4, 76.
 Distribution, 38, 39, 166.
 Disutility, 60, 121.
 Division of labour, 17-22, 64, 123-4, 154-7, 181, 204.
 Douglas, P. H., 174.
 Duopoly, 85-7, 263.
- Economic development, 205-8.
 — goods, 50-1.
 — man, 6, 24, 117.
 — motive, 5.
 Economics and politics, 246, 260-1, 268.
 — and technology, 115-16, 143-4.
 — classical, 4, 5, 38, 198, 207, 228-30, 264.
 — modern, 39, 200-1, 205-6, 208, 230, 264-7.
 — teleological, 267-8.
 Engel, Ernst, 221-2.
 Entrepreneur, 127-35, 186, 190-1, 193-4, 195-201, 223, 227-33, 251.
 Equilibrium, 78, 83-4, 85-7, 110-14, 130-1, 154, 201, 202-3, 263.
 Exchange, 21-3, 38-9, 61-2.
 — economy, 122-6, 150, 238, 245, 250.
 — indirect, 87-9, 29-3.
 — isolated, 67-9.
 — laws of, 100-6, 111, 174, 187.
 — medium of, 88-90.
 — theory of, 40-1, Part Two, Chapter Two.
- Factors of Production, 116-17.
 — — exchange of, 122-4.
 — — imperfect divisibility of, 154-7, 181, 192-5.
 — — prices of, 133-7.
 — — transference of, 243.
 Fisher, I., 106.
 Fluctuations, 225, 244, 247-61.
- Giffen, Sir R., 110.
 Gold, 90-1.
 Goods, 4-7, 38-9, 47-51.
 — capital, 187, 256-9.
 — competing and completing, 106.
 — consumption, 117-18, 255-60.
 — intermediate, 187.
 — production, 187, 252, 256-9.
 Gossen, H. H., 56.
 — First Law of, 45-7, 54-5, 154, 165.
 — Second Law of, 58-9, 61-2, 111, 120, 135, 165.
- Hicks, J. R., 66, 99, 107, 234, 235.
- Impartiality, 30-2, 264-8.
 Income, 192, 196, 206, 221, 222.
 Increasing returns, 152-61.
 Indeterminacy, 68, 176.
 Index numbers, 171.
 Inequality, 26.
 Inheritance, 26.
 Interest, 190-2, 196, 251-2, 258.
 International trade, 141-3.
 Intervention, 205, 241-6.
 Invention, 206, 227-30, 234-7.
 Investment, 187-90, 251.
 Isolated economy, 42, 59-62, 119-22, 150.
- Jevons, W. St., 254.
 Joint products, 143-9.
- Knight, F. H., 43, 201.
- Labour, 59-60, 116, 120-1, 151-2, 212, 226, 229-33.
 — demand for, 172-9, 227-34.
 — division of, 17-22, 64, 123-4, 154-7, 181, 204.
 — legislation, 242-3.
 — marginal product of, 176, 234-7.
 — redistribution of, 239-40, 243.

- Labour, supply of, 124-6, 173-6, 217.
Laissez-faire, 260.
 Land, 116, 151-2, 166-9, 212, 226.
 Leisure, 121, 233.
 Luddites, 228.
- Machine production, 19, 20, 184, 188-90, 227-33.
 Malthus, T. R., 210-17.
 Marginal acre, 168.
 — buyer, 74, 78, 81, 163-4.
 — cost, 137-9, 142, 145, 146.
 — pairs, 78-9.
 — producers, 138-40.
 — product, 164, 167, 176, 233-4.
 — productivity, 164-5, 191, 194-5.
 — rate of substitution, 66-7.
 — seller, 79.
 — utility, 55-7, 61, 64-7, 72-3, 91, 92-6, 100-1, 106-8, 135.
 Market, 14, 22-4, 27, 110-14, 155-6, 163, 181, 229, 237-8, 240, 263.
 Marshall, A., 146.
 Marx, K., 199-200, 207-8, 266-7.
 Mass production, 152-3, 155, 157, 224.
 Mill, John Stuart, 36.
 Mobility, 26-8, 136, 142-3, 278-9, 239-40.
 Money, 23-4, 90-1, 187, 246.
 — and the trade cycle, 256-9.
 — demand for, 98-100, 263.
 — quantity theory of, 97-8.
 — value of, 96-100.
 Monopoly, 73-6, 82, 84, 86, 104-6, 148, 158-60, 163-4, 175, 196-7, 206, 240, 242, 263.
 — bilateral, 69, 176.
- Net advantages, 177-9.
 Non-competing groups, 142-3, 178-9, 195.
 Oligopoly, 85, 263.
- Pigou, A. C., 86, 176.
 Policy, 208-9, 237-46, 259-61, 264-5.
 Politics and economics, 246, 260-1, 268.
 Population, 201, 203, 210-17, 219-20.
 Price, 94, 100-6, 175, 223.
 — and cost, 128-33.
 — — — in perfect competition, 161-2.
 — — — in imperfect competition, 162-4, 244.
 — in the trade cycle, 251-3, 255-6.
 — interdependence, 117-18, 127-8, 202-5, 225-7.
 — of capital, 190-2.
 Private enterprise, 25, 124, 186, 189-90, 198-9, 200.
 Private property, 21, 22, 245.
 Producer's surplus, 135-43, 196-7.
 Production, 38-41.
 — complex, 149-66, 200.
 — co-operative, 123-4.
 — in the Trade Cycle, 251-3
 — goods, 48-9, 255-6.
 — single-factor, 126-35.
 — theory of, Part Two, Chapter Three.
 Profits, 195-201, 226, 229, 248, 251.
- Rent, 166-9, 196-7, 221-2.
 Ricardo, D., 167, 229.
 'Rigging', 240.
 Risk, 198-9, 201, 251.
 Robbins, L., 115.
 Roundabout production, 179-87.
- Sales-cost, 163.
 Satisfaction, 7-9, 58-62, 65, 111, 121-2, 220-1.
 Saving, 187-90, 192, 204, 206, 251, 263.
 Say, J. B., 229.
 Scarcity, 11-13.

- Schumpeter, J., 207.
 Schwabe's law, 221-2.
 Scientific method, 29-37.
 — laws, 265-6.
 Self-interest, 5.
 Slump, 248, 254, 259-60.
 Smith, Adam, 19, 155, 177, 229.
 Social data, 16-17, 186, 218-19, 238-9.
 Sraffa, P., 160, 163.
 Standard of living, 45, 232.
 State, 241-6, 260-1.
 Stationary economy, 201.
 Strigl, R. v., 127.
 Subsistence means, 183-7, 211-14.
 — minimum of, 174-6.
 Substitutes, 96, 106-10, 159.
 Substitution, Marginal rate of, 66-7.
 Supply, 71-2, 94, 101, 124-6, 202-5.
 — joint, 143-9, 251.
 — of labour, 124-6, 173-6, 217.
 — of capital, 191-2.
 Surplus, Consumer's, 79-81.
 — producer's, 135-43, 196-7.
 — product, 199.
 — value, 199.
 Tastes, 201, 203, 217-27.
 Technical coefficient, 119-20, 135-44, 197, 204.
 Technique, 201, 203, 216-17, 219-21, 223, 227-37.
 Technology and economics, 115-16, 143-4.
 Time preference, 192, 251.
 Trade cycle, 247-61, 263.
 Transference, 237-40, 243.
 Truck system, 171.
 Turgot, 151.
 Uncertainty, 99-100, 198-9, 201, 263.
 Utility, 51-7, 267.
 — diminishing, 54-7, 154.
 — marginal, 55-7, 61, 64-7, 72-3, 91, 92-6, 100-1, 106-8, 135.
 — measurability of, 60-1, 65-7, 80, 92-3.
 Value, 71, 92-3, 133, 266.
 — of money, 96-100.
 Wages, 69, 128, 169-70, 195, 217, 231.
 — differences, 176-7.
 — real and money, 170-2.
 — system, 169-70.
 — theory of, 167, 169-79.
 Wage-earner, 24-5, 200, 228, 241.
 Wants, 7-9, 42-7, 220-1, 232.
 — marginal, 56.
 — *Minimum Sensibile* of, 47, 154.
 Wealth, 51, 221.
 West, Sir E., 151.
 Wicksell, K., 230.
 Wicksteed, P. H., 65, 93, 105, 133.